# 線形代数

## 基礎と応用

新井仁之 著

日本評論社

# はじめに—線形代数をなぜ学ぶのか

大学に入学して最初に学ぶ数学は，古くから解析学と線形代数ということになっている．大学によっては解析学の代わりに微分積分という科目名が用いられることもあるが，ここでは解析学ということにする．ところで著者は以前から，最初に学ぶ数学が，今日でも解析学と線形代数でよいのだろうかということを考えていた．確かに数学科で本格的に数学を学ぶ場合，この二つの分野はすべての基礎となっている．しかしより広く考えて，工学系，理学系あるいは数学を必要とする文系の学生にとって，ほかのより良い選択肢はないのだろうか？

結論を先に言えば，歴史的ないきさつはともかく，今日の科学技術の状況を考慮すれば，解析学と線形代数という選択は結果として非常によいものであったと著者は考えている．読者が線形代数を学ぶ際の参考にもなると思うので，本文に入る前にその理由を述べておきたい．

まず解析学についてであるが，科学技術に関連した分野では，多くの関係者が言うように微分方程式の解析，フーリエ解析，確率論などを必要とする場面に頻繁に遭遇する．したがって科学技術に深くかかわるためには，これらの基礎をなす解析学の知識が必要不可欠であることは衆目の一致するところであろう．

次に解析学以外の科目で必要度の大きいものは何かを考えてみよう．現代社会では，科学技術はもちろんのこと，多くの分野で，さまざまなデータを解析することが必要になる．扱うデータの個数は，たとえばアナログ信号のように非可算無限個の場合もあるが，現実的には有限個の場合が多い．またかなりの頻度でコンピュータを使った解析が行われるが，しかしコンピュータが扱えるデータの個数は有限個である．そのため，有限個のデータを扱う数学の重要度が今日では急速に増してきている．現代では分野にもよるが，有限データに関する数学のかなり高度な素養を身に着けておくことが要求されている．たとえば，情報処理，画

像解析，時系列，統計学などを扱う分野などがその例であろう．

このような有限データを扱う数学は何か？　じつはその一つが線形代数なのである．著者は次のように考えている．

**無限個のデータを扱う数学が解析学であり，有限個のデータを扱う数学が線形代数である．**

このような理由から，これからの社会で生活していく大学生諸君が，大学で線形代数と解析学を学んでおくことは，たいへん望ましいことであるといえよう．

さて，それでは学生諸君は線形代数の何を学んでおけばよいのか，そしてそれを学んでいるとどのようなメリットがあるのだろうか．このことは，これから線形代数を学ぶ学生諸君にとっては自然な質問であり，一方，大学教師にとっては講義の準備をする際に常に自問自答しておかねばならない教育上の問題である．この問題に対する著者の答えを記したものが本書である．

本書は「線形代数は有限個のデータを扱う数学である」という視点に立脚して，線形代数の入門的な事柄からはじめ，いくつかの有用な応用例あるいは応用の原理が書いてある．たとえば有限個のデータを扱う数学の典型的な例として，多変量解析，離散フーリエ解析，離散ウェーブレット解析などがあるが，本書ではその線形代数的な側面を大きく取り上げている．また近年，実用的な重要性を増しつつある一般逆行列と特異値分解についても，入門的事項を詳しく解説した．これらをとおして，読者は線形代数の有用性を十分理解いただけると思う．

本書の概略を述べておく．この本は四部構成になっている．第 1 部は行列と行列式について，第 2 部は線形構造と基底，第 3 部は各種の応用，そして第 4 部は線形代数の抽象化について解説してある．第 1 部は初学者が無理なく線形代数の世界に入れるよう，行列と連立方程式について解説した．第 2 部は第 1 部より理論的色彩が濃く，数ベクトル空間とその線形部分空間に関わる事項が解説されている．とりわけ「線形部分空間」，「基底」，「正規直交基底」およびその役割が詳しく述べられている．第 2 部は数学的な議論が積み重ねられていくが，組み立てられた理論は第 3 部で展開される応用の根幹をなすことになる．第 3 部では，多変量解析，離散フーリエ解析，離散ウェーブレット解析など実用的な数学が，その基礎的な部分で本質的に線形代数を用いていることを記した．またこれらに関連して，特異値分解，一般逆行列の章も設けた．第 4 部では線形代数の抽象化が

おこなわれる．抽象化することにより，数ベクトル空間からテンソル空間，関数空間へと扱う対象が広がっていく．その結果応用の範囲も拡大されていく．線形空間はさらに無限次元の線形位相空間へと抽象化されていくが，それは本書では扱わない．

なお，各章の最初の部分に「ポイント」という段落を設け，その章で何を理解してほしいかを要約してある．学習の際の参考にしてほしい．

この本の構想は拙著『フーリエ解析学』(朝倉書店，2003 年刊) の執筆中に練り始めた．『フーリエ解析学』では，まず有限個の点を扱う離散フーリエ解析により調和解析の構造を浮き彫りにし，それから無限を扱う場合へと深めていくというストーリーを書いた．その際に離散フーリエ解析は線形代数の特別な場合であるということを痛感した．この経験が本書執筆の直接の契機となった．

本書の出版にご尽力いただいた亀書房の亀井哲治郎氏，亀井英子氏に感謝したい．

なお図の多くは新井しのぶによって描かれたものである．数値計算，グラフ描画には MATLAB® を，図の作成には Adobe® Illustrator® を用いた．

本書に関する情報を著者のホームページ

http://www4.ocn.ne.jp/~arai

に適宜掲載していくので，本書と併せて参照いただきたい．

2005 年 7 月

新井仁之

本書では次の記号を用いている．

1. 定理の証明の後に ■ の記号を付しているが，これは証明が終了したことを表す．

2. 複素数 $z = a + \sqrt{-1}\,b$, ただし $a, b$ は実数とするとき，$\overline{z} = a - \sqrt{-1}\,b$ とする．$\overline{z}$ を $z$ の複素共役という．また $|z| = (a^2 + b^2)^{\frac{1}{2}}$ を $z$ の絶対値という．特に $z = a$ の場合は明らかに $\overline{z} = a = z$ である．

3. 二つの命題 (あるいは二つの式)$P$ および $Q$ に対して，$P$ が成り立つための必要十分条件は $Q$ が成り立つことであることを $P \Longleftrightarrow Q$ と表す．

4. 集合 $A, B$ に対して，$A \subset B$ は $A \subsetneqq B$ ではなく $A \subseteq B$ を意味するものとする．

5. 用語を定義する際に，「本書では*** という」という言い方をしている部分がある．これは便宜上本書での呼称を定めているもので，一般的なものとは限らないものもある．

# 目　次

はじめに—線形代数をなぜ学ぶのか　i

## 第1部　基礎編——行列と行列式——

**第1章　数ベクトル空間，線形写像，基底　3**

1.1　数ベクトル …… 3
1.2　数ベクトルの算術 (線形演算) …… 7
1.3　線形写像 …… 10
1.3.1　線形写像の例 …… 12
1.3.2　線形写像と連立 1 次方程式 …… 16

**第2章　行列と行列の演算　18**

2.1　行列 …… 18
2.2　行列の演算 …… 20
2.2.1　行列の和と差 …… 20
2.2.2　行列の定数倍 …… 21
2.2.3　行列の積 …… 22
2.2.4　行列の積と線形写像 …… 24
2.2.5　転置行列と共役行列 …… 26
2.2.6　行列のアダマール積 …… 27
2.3　基本的な行列の例 …… 27
2.4　逆行列 …… 31

**第3章　線形写像と行列　37**

3.1 線形写像の行列による表現 …… 37
3.2 行列と線形写像の演算 …… 40
3.3 写像と逆写像 …… 42

**第 4 章 ガウスの消去法 47**
4.1 具体例 …… 48
4.2 より一般の場合のガウスの消去法 …… 52
4.2.1 上三角行列への変形 …… 53
4.2.2 前進代入法と後退代入法 …… 55
4.3 基本行列の積による行列の変形 …… 56
4.3.1 具体例 …… 57
4.3.2 一般の場合 …… 60
4.4 逆行列の計算——掃き出し法 …… 64

**第 5 章 行列式 70**
5.1 置換 …… 71
5.2 置換の符号と偶置換，奇置換 …… 75
5.3 行列式 …… 79
5.4 行列式の基本的な性質 …… 84
5.4.1 行に関する変形 …… 84
5.4.2 列に関する変形 …… 89
5.4.3 行列式のベクトルによる表示 …… 91

**第 6 章 行列式の余因子展開とその応用 95**
6.1 余因子展開 …… 95
6.2 余因子展開を用いた行列式の計算 …… 98
6.3 行列式と余因子を用いた逆行列の計算 …… 100
6.4 行列式と連立 1 次方程式の解法 …… 102

**第 7 章 いろいろな行列の行列式 106**
7.1 ファンデルモンド行列式と補間多項式への応用 …… 106
7.2 置換行列 …… 111
7.3 巡回行列の行列式 …… 115
7.4 固有多項式 …… 117

7.5 小行列と小行列式 …… 118

**第 8 章 ブロック行列 124**

8.1 ブロック行列の演算 …… 124
8.2 ブロック行列の逆行列 …… 129
8.3 ブロック行列の行列式 …… 132

## 第 2 部 理論編 —— 線形構造と基底 ——

**第 9 章 基底と部分空間 137**

9.1 線形独立，線形従属 …… 137
9.2 線形包 …… 140
9.3 線形部分空間とその基底 …… 144

**第 10 章 内積と正規直交基底 154**

10.1 内積と直交性 …… 154
10.2 正規直交基底 …… 160
10.3 シュミットの直交化法 …… 162
10.4 直交射影と直交補空間 …… 164
10.5 最良近似への応用 …… 172
10.6 線形写像の値域と核 …… 176

**第 11 章 行列の階数 180**

11.1 一般論 …… 180
11.2 階数の計算とピボット …… 186
11.3 行列の標準化 …… 190

**第 12 章 連立 1 次方程式の一般解 193**

12.1 解の存在と一般解 …… 193
12.2 連立 1 次方程式の一般解の求め方 …… 198

**第 13 章 基底変換と行列の対角化 207**

13.1 基底変換 …… 210
13.2 対角化と固有値 …… 215

13.3 正規直交基底による対角化 …… 227
13.4 エルミート形式とクーラン-フィッシャーの定理 …… 238
13.5 幾何的な問題と主成分分析への応用 …… 244

**第 14 章 行列の分解定理** **251**
14.1 $LU$ 分解 …… 251
14.2 $LDM^*$ 分解 …… 260
14.3 コレスキー分解 …… 263
14.4 $QR$ 分解 …… 265
14.4.1 ハウスホルダー $QR$ 分解 …… 265
14.4.2 シュミットの方法による $QR$ 分解 …… 269

# 第 3 部 応用編

**第 15 章 一般逆行列とその応用** **273**
15.1 一般逆行列 …… 273
15.2 ムーア-ペンローズ一般逆行列 …… 279
15.3 連立方程式の最小 2 乗解への応用 …… 285
15.4 データの直線，曲線による当てはめへの応用 …… 288
15.5 種々の一般逆行列 …… 292
15.5.1 反射型一般逆行列 …… 293
15.5.2 最小 2 乗型一般逆行列 …… 294
15.5.3 ノルム最小型一般逆行列 …… 295

**第 16 章 特異値分解とその応用** **297**
16.1 行列の特異値分解 …… 297
16.2 特異値標準形と一般逆行列 …… 305
16.3 特異値分解と最小 2 乗解 …… 309
16.4 低階数の行列による近似とディジタル画像 …… 311

**第 17 章 多変量解析と線形代数** **319**
17.1 いくつかの基本概念 …… 319
17.2 回帰分析 …… 324
17.3 主成分分析 …… 326

**第 18 章 離散フーリエ解析への応用 332**

18.1 フーリエ解析とは何か …… 333
18.2 フーリエ基底 …… 335
18.3 フィルタリングとその応用 (ノイズ除去) …… 340
18.4 循環相関積 …… 345
18.5 フーリエ行列と巡回行列 …… 348
18.6 スペクトログラム …… 352

**第 19 章 離散ウェーブレットへの応用 356**

19.1 準備 …… 357
19.2 サブバンド・フィルタ・バンク …… 365
19.3 2 チャネル最大間引きフィルタ・バンク …… 371
19.4 多重解像度近似 …… 375
19.4.1 一般化多重解像度近似 …… 375
19.4.2 最大間引きフィルタ・バンクと多重解像度解析 …… 378
19.5 ウェーブレットの例 …… 382
19.5.1 ハール・ウェーブレット …… 382
19.5.2 ドブシー・ウェーブレット …… 388
19.5.3 3 チャネル 2 間引きフィルタ・バンクの例 …… 389
19.6 ウェーブレットの応用例 (特異性の検出) …… 390

**第 20 章 整数値行列とその応用 392**

20.1 スミス標準形 …… 392
20.2 整数値行列による格子の生成 …… 399

# 第 4 部 線形代数の抽象化

**第 21 章 線形空間 407**

21.1 線形空間の定義と例 …… 407
21.2 線形写像と行列 …… 416
21.3 座標変換について …… 418
21.4 内積 …… 421

**第 22 章　テンソル積と外積** **427**
22.1　線形空間のテンソル積 ………… 427
22.2　線形写像のテンソル積 ………… 434
22.3　画像処理と線形写像のテンソル積 ………… 441
22.4　反変テンソル，共変テンソル ………… 444
22.5　交代テンソル ………… 446
22.6　テンソル代数と外積代数 ………… 448

**第 23 章　$k$-ベクトルと $k$-形式** **455**
23.1　$k$-ベクトルと線形空間の向き ………… 455
23.2　$k$-形式 ………… 460
23.3　$k$-ベクトルに対する内積 ………… 465
23.4　$k$-形式に対する内積 ………… 470

# 付録

**付録 A　置換を互換の積に分解する方法** **479**

**付録 B　行列式の幾何学的意味** **481**

**付録 C　行列に対するノルム** **483**

**付録 D　ジョルダン標準形** **487**

**付録 E　問題の解答** **492**

**参考文献** **530**

# 第１部
# 基礎編
## ——行列と行列式——

# 第1章　数ベクトル空間，線形写像，基底

わたしたちの身の回りには，多くの「変換」が存在する．たとえば，携帯電話を例にすれば，まず送り手は電話機により音声あるいは画像をディジタル信号へ変換し，それを伝送して他人に伝える．一方，受け手はそのディジタル信号を再び音声，画像に変換して理解する．さらに伝送する際にも，単にディジタル信号に変換するだけでなく，データは適切な量に圧縮された上で伝送される．この圧縮も変換の一つである．これはほんの一例にすぎないが，わたしたちはほかにも多くの「変換」に取り囲まれている．これらの変換の中で，基本的でありかつ重要なものが線形変換，別名線形写像と呼ばれているものである．本章では数ベクトルと線形写像の定義と例を述べる．これらの概念は線形代数の出発点をなすものである．

本書では実数全体のなす集合を $\boldsymbol{R}$ で表し，複素数全体のなす集合を $\boldsymbol{C}$ により表す．また，$\boldsymbol{K}$ により $\boldsymbol{R}$ または $\boldsymbol{C}$ を表すことにする．$\boldsymbol{Z}$ により整数全体を表す．

**ポイント1**　本章では数ベクトル空間，線形写像，基底の定義をおさえておいてほしい．また線形写像と連立1次方程式との関連も重要事項である．

## 1.1　数ベクトル

データは数の集まりであるが，それを表現するために数ベクトルというものを導入する．$N$ 個の数をたてに並べたもの

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

を $N$ **次元数ベクトル**あるいは $N$ **次元縦ベクトル**という．そしてこのベクトルを構成する数 $x_1, \cdots, x_N$ を $N$ 次元数ベクトルの**成分**といい，$x_j$ をこの数ベクトルの**第** $j$ **成分**という．特に各 $x_j$ が実数からなるものを $N$ **次元実ベクトル**，また各 $x_j$ が複素数からなるものを $N$ **次元複素ベクトル**という．$N$ **次元数ベクトル**を略して**数ベクトル**あるいは**縦ベクトル**ということもある．ベクトルに対して，実数あるいは複素数を**スカラー**という．本書では，スカラーと数ベクトルを区別するため，$N$ 次元数ベクトルは $\mathbf{x}, \mathbf{y}, \mathbf{z}, \cdots$ のように太字で表す．また記号を簡略にするため，

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

を単に $\mathbf{x} = (x_j)_{j=1,\cdots,N}$，あるいは混乱のない場合はさらに $\mathbf{x} = (x_j)$ と略記する．

二つの $N$ 次元数ベクトル $\mathbf{x} = (x_j)_{j=1,\cdots,N}$，$\mathbf{y} = (y_j)_{j=1,\cdots,N}$ に対して，

$$x_1 = y_1, \quad \cdots, \quad x_N = y_N$$

であるとき，この二つの数ベクトルは等しいといい，

$$\mathbf{x} = \mathbf{y}$$

と表す．

このような数ベクトルで表されるものは非常に多い．典型的な例をいくつかあげておこう．

### 数値データ，音声データ

私たちはしばしば有限個の数値からなるデータを扱う．たとえば，日々の気温の変化，太陽の黒点の数をはじめ，時刻とともに変化する自然現象を観測し，数値として記述したものなどである．また音声や画像を数値データとして扱うこともできる．音声を有限個の数値にするには次のようにすればよい．音声は空気圧の変動によって生じるので，縦軸が空気圧の変化，横軸が時間を表すようにすれば，たとえば図 1.1 の上図のような波形となる．これは筆者の「アー」と言った音をおよそ 0.028 秒間分切り取ったものである[1)]．

---

1) 正確には，IC レコーダーで録音したもので，44.1k[Hz] でサンプリングされたものを連続的なグラフにしたものである．

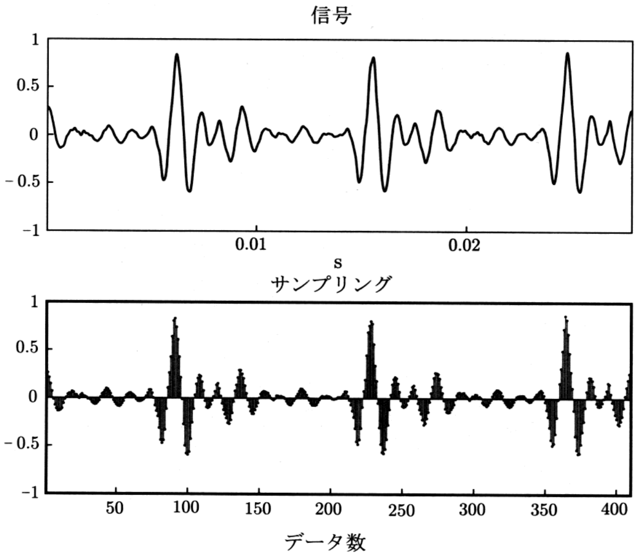



図 1.1

一方，図 1.1 の下図は $[0, 0.028]$ を 409 等分した時刻 $0 = t_1 < t_2 < \cdots < t_{410}$ での値 $x(t_1), x(t_2), \cdots, x(t_{410})$ をサンプルとしてとったものである．$x_k = x(t_k)\,(k = 1, \cdots, 410)$ と表し，それを

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{410} \end{pmatrix}$$

とすると音声信号をたとえば 410 次元実ベクトル $\mathbf{x}$ として扱うことができる．

### ディジタル画像

画像も有限個の数ベクトルで表すことができる．実際，ディジタル画像などは有限個の整数からなる数ベクトルとして表されている．たとえば図 1.2 のディジタル写真は $256 \times 256$ 画素 (ピクセル) で 256 階調のものである．

256 階調とは，各画素の濃淡が黒 → 灰色 → 白と連続しているのを 0 から 255 の整数で表している．したがって，このディジタル写真は 0 から 255 の間の整数を $256 \times 256 = 65536$ 個並べたもの



図 1.2　256 × 256 ピクセルの 256 階調のディジタル画像

$$\begin{matrix} x_{1,1} & \cdots & x_{256,1} \\ \vdots & \ddots & \vdots \\ x_{1,256} & \cdots & x_{256,256} \end{matrix}$$

として表されている．これを一列に並べ替えて

$$\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{1,256} \\ \vdots \\ x_{256,1} \\ \vdots \\ x_{256,256} \end{pmatrix}$$

とすれば，ディジタル画像は 65536 次元実ベクトルとなっている．なお画像は数ベクトルの**テンソル積**(テンソル積は第 22 章で学ぶ) とみなすほうが扱いやすいことが多い．

このように $N$ 個の実数の組で表されるデータを $N$ 次元数ベクトルと考えて，各種データの処理に線形代数を利用することができる．

**幾何学における座標とベクトル**

平面上の点の位置は直交座標を使えば2次元実ベクトル $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ により表すことができる．また空間ならば3次元実ベクトル $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ によって表すことができる．

もう一つの典型的な例として，高等学校で学んだ「ベクトル」がある．「ベクトル」とは向きと大きさを持った量である．たとえば平面ベクトル $\overrightarrow{a}$ は座標平面の基本ベクトル $\overrightarrow{e_1}$ と $\overrightarrow{e_2}$ により $\overrightarrow{a} = a_1\overrightarrow{e_1} + a_2\overrightarrow{e_2}$ と表せるので，$\overrightarrow{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ と成分表示できる．この成分表示により平面ベクトルは2次元実ベクトル $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ によって表せたことになる．また空間ベクトルはこのような成分表示を用いれば3次元実ベクトルで表せる．

## 1.2　数ベクトルの算術 (線形演算)

前節で述べたような数値データや幾何的なベクトルを数学的に解析するために，数ベクトルに関するいくつかの演算を導入する．すでに述べたように $\boldsymbol{K}$ により $\boldsymbol{R}$ または $\boldsymbol{C}$ を表す．どちらを表すかはあらかじめ定めておく必要があるが，どちらでも成り立つ場合は特に指定せずに議論を進める．

以下，本書では

$$\boldsymbol{K}^N = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} : x_1, \cdots, x_N \in \boldsymbol{K} \right\}$$

とする．特に $\boldsymbol{K} = \boldsymbol{R}$ とする場合は $\boldsymbol{K}^N$ を $\boldsymbol{R}^N$，また $\boldsymbol{K} = \boldsymbol{C}$ とする場合は $\boldsymbol{K}^N$ を $\boldsymbol{C}^N$ と表す．$\boldsymbol{K}^N$ を $N$ **次元数ベクトル空間**という．

二つの $N$ 次元数ベクトル

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \in \boldsymbol{K}^N$$

に対してその**和**を

$$\mathbf{x}+\mathbf{y}=\begin{pmatrix} x_1+y_1 \\ \vdots \\ x_N+y_N \end{pmatrix}$$

と定義する．明らかに $\mathbf{x}+\mathbf{y}\in \boldsymbol{K}^N$ である．$\mathbf{x}+\mathbf{y}$ を $\mathbf{x}$ と $\mathbf{y}$ との**重ね合わせ**ともいう．また，スカラー $\alpha\in\boldsymbol{K}$ と $N$ 次元数ベクトル $\mathbf{x}\in\boldsymbol{K}^N$ に対してその積 $\alpha\mathbf{x}$ を

$$\alpha\mathbf{x}=\begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_N \end{pmatrix}$$

と定義する．この演算をベクトルの**スカラー倍**という．$\alpha\mathbf{x}\in\boldsymbol{K}^N$ である．これらの和とスカラー倍を総称して**線形演算**という．

**注意 1.1**　もし $\boldsymbol{K}=\boldsymbol{R}$ とあらかじめ指定した場合には，$\alpha\in\boldsymbol{K}=\boldsymbol{R}$ と $\mathbf{x}\in\boldsymbol{K}^N=\boldsymbol{R}^N$ に対して $\alpha\mathbf{x}$ が定義されている．実際 $\alpha\mathbf{x}\in\boldsymbol{R}^N$ である．しかしこの場合はスカラーとして複素数を考えない．なぜならば $\alpha\in\boldsymbol{C}$ のときは一般に $\alpha\mathbf{x}\notin\boldsymbol{R}^N$ だからである．ただし $\boldsymbol{K}=\boldsymbol{C}$ と指定した場合には $\alpha\in\boldsymbol{K}=\boldsymbol{C}$ と $\mathbf{x}\in\boldsymbol{K}^N=\boldsymbol{C}^N$ に対して $\alpha\mathbf{x}$ が定義され，$\alpha\mathbf{x}\in\boldsymbol{C}^N$ である．$\boldsymbol{K}$ として $\boldsymbol{R}$ と指定するかあるいは $\boldsymbol{C}$ と指定するかによって，スカラーは実数か複素数かの違いがでてくる．

一般に $n$ 個の $N$ 次元数ベクトルとスカラー

$$\mathbf{x}_1=\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{N,1} \end{pmatrix},\quad \cdots,\quad \mathbf{x}_n=\begin{pmatrix} x_{1,n} \\ \vdots \\ x_{N,n} \end{pmatrix}\in\boldsymbol{K}^N,$$

$$\alpha_1,\cdots,\alpha_n\in\boldsymbol{K}$$

に対して

$$\alpha_1\mathbf{x}_1+\cdots+\alpha_n\mathbf{x}_n=\begin{pmatrix} \alpha_1 x_{1,1}+\cdots+\alpha_n x_{1,n} \\ \vdots \\ \alpha_1 x_{N,1}+\cdots+\alpha_n x_{N,n} \end{pmatrix}$$

と定義する．$\alpha_1\mathbf{x}_1+\cdots+\alpha_n\mathbf{x}_n$ は $\sum$ の記号を用いて

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n = \sum_{j=1}^{n} \alpha_j \mathbf{x}_j$$

とも表す．$\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$ の形の和を $\mathbf{x}_1, \cdots, \mathbf{x}_n$ の**線形結合**あるいは **1 次結合**という．

後に用いる基本的な用語を定めておきたい．$N$ 次元数ベクトルで，すべての成分が 0 になっているものを $N$ **次元零ベクトル**といい，$\mathbf{0}_N$ または単に $\mathbf{0}$ で表す．また $N$ 次元数ベクトル

$$\mathbf{e}_1^{(N)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2^{(N)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad \mathbf{e}_N^{(N)} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{1.1}$$

を $N$ **次元基本ベクトル**という．特に混乱のない場合は $\mathbf{e}_j^{(N)} = \mathbf{e}_j$ と略記する．明らかに任意の $N$ 次元数ベクトル $\mathbf{x}$ は

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = x_1\mathbf{e}_1^{(N)} + \cdots + x_N\mathbf{e}_N^{(N)} \tag{1.2}$$

と表せる．

$N$ 個の数 $x_1, \cdots, x_N \in \boldsymbol{K}$ を横に並べたもの

$$(x_1, \cdots, x_N)$$

を $N$ **次元横ベクトル**あるいは単に**横ベクトル**という．$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x}^T = (x_1, \cdots, x_N)$$

と表し，$\mathbf{x}^T$ を $\mathbf{x}$ の転置という．なお本によっては ${}^t\mathbf{x}$ と表しているものもある．

## 1.3　線形写像

線形代数の主な対象の一つは，ベクトルを別のベクトルに写す線形写像と呼ばれる変換である．線形写像は幾何学や解析で基本的な役割を果たす．

まず一般的な写像の定義からはじめよう．$N$ 次元数ベクトルに対して，ある規則で $M$ 次元数ベクトルを定める対応 $T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への**写像**という．写像 $T$ によって $\mathbf{x} \in \boldsymbol{K}^N$ が $\mathbf{a} \in \boldsymbol{K}^M$ に写されることを $T(\mathbf{x}) = \mathbf{a}$ と表す．そしてこの写像を $T$ と表す．

写像 $T$ が $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への写像であることを

$$T : \boldsymbol{K}^N \to \boldsymbol{K}^M$$

という記号で表す．今後 $T(\mathbf{x})$ を単に $T\mathbf{x} = \mathbf{a}$ と表すこともある．

**例 1.1**　たとえば $\mathbf{x} = (x_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ に対して，

$$T(\mathbf{x}) = \begin{pmatrix} \dfrac{x_1 + x_2}{2} \\ \dfrac{x_2 + x_3}{2} \\ \vdots \\ \dfrac{x_{N-1} + x_N}{2} \end{pmatrix} \tag{1.3}$$

とすると，$T$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^{N-1}$ への写像である．また

$$T(\mathbf{x}) = \frac{x_1 + \cdots + x_N}{N} \tag{1.4}$$

とすると，$T$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^1$ への写像である．

写像は図 1.3 のように表されることがある．図 1.3 は写像 $T$ をある装置にたと

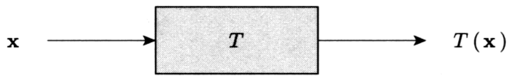



図 1.3

えているといえよう．つまり，$\mathbf{x}$ という $N$ 次元数ベクトルを $T$ に入力すると，それが一定の規則に基づいて $T(\mathbf{x})$ となって出力されると考えるのである．

写像の中で，本書で扱うのは主に次に定義される「線形写像」と呼ばれるものである．

**定義 1.1**　$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への写像 $T$ が，$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$，$\alpha \in \boldsymbol{K}$ に対して

$$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}), \tag{1.5}$$

$$T(\alpha\mathbf{x}) = \alpha T(\mathbf{x}) \tag{1.6}$$

をみたすとき，$T$ は**線形**である，あるいは $T$ は**線形写像である**という．

線形とは，$\mathbf{x} \in \boldsymbol{K}^N$ に対する出力が $T(\mathbf{x})$ であり，$\mathbf{y} \in \boldsymbol{K}^N$ に対する出力が $T(\mathbf{y})$ であるとき，入力 $\mathbf{x}+\mathbf{y}$ に対して $T(\mathbf{x})$ と $T(\mathbf{y})$ を重ね合わせたものが出力され，さらに $\alpha\mathbf{x}$ を入力すると $T(\mathbf{x})$ の $\alpha$ 倍したものが出力されることである (図 1.4 参照).

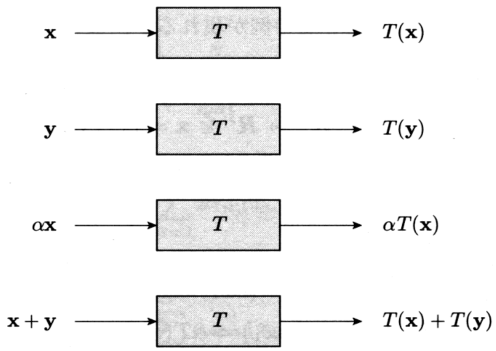



図 1.4

線形写像は次に示すように零ベクトルを必ず零ベクトルに写す．

**命題 1.1**　$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像とする．このとき，$T(\mathbf{0}_N) = \mathbf{0}_M$ である．

**証明.**　$T(\mathbf{0}_N) + T(\mathbf{0}_N) = T(\mathbf{0}_N + \mathbf{0}_N) = T(\mathbf{0}_N)$ であるから，$T(\mathbf{0}_N) = T(\mathbf{0}_N) - T(\mathbf{0}_N) = \mathbf{0}_M$．■

**注意 1.2**　$\boldsymbol{K}$ はあらかじめ $\boldsymbol{R}$ または $\boldsymbol{C}$ であることを指定されているが，$\boldsymbol{K} = \boldsymbol{R}$ と指定されている場合，$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像は $\boldsymbol{R}^N$ から $\boldsymbol{R}^M$ への線形写像を意味している．この場合は，線形写像の定義 (1.6) におけるスカラー $\alpha$ としては $\boldsymbol{R}$ の要素しか考えない．すなわち $\boldsymbol{R}^N$ から $\boldsymbol{R}^M$ への線形写像 $T$ に対しては $\alpha \in \boldsymbol{R}$ ならば $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$ が成り立つが，$\alpha \in \boldsymbol{C}$ に対しては $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$ が成り立つことは保障されていないのである．しかし $\boldsymbol{K} = \boldsymbol{C}$ の場合，$\boldsymbol{K}^N$ から

$\boldsymbol{K}^M$ への線形写像 $T$ に対しては $\alpha \in \boldsymbol{C}$ について $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$ が成り立つ.

**注意 1.3**　本書では扱わないが，線形とは限らない写像を**非線形写像**という．たとえば

$$T(\mathbf{x}) = \begin{pmatrix} |x_1| \\ \vdots \\ |x_N| \end{pmatrix}$$

は非線形写像である.

### 1.3.1　線形写像の例

本書を通してさまざまな線形写像の例が現れるが，ここでは線形写像の簡単な例と重要な例をいくつかあげておく.

**例 1.2**　$a$ を実数とする．$T : \boldsymbol{R}^1 \to \boldsymbol{R}^1$ を $\mathbf{x} = (x_1) \in \boldsymbol{R}^1$ に対して，$T(\mathbf{x}) = (ax_1)$ とすると，明らかに $T$ は線形である．$b$ を $0$ でない実数とし，$T(\mathbf{x}) = (ax_1 + b)$ とすると，$T$ は線形ではない.

**証明.** $\mathbf{x}, \mathbf{y} \in \boldsymbol{R}^1$, $\alpha \in \boldsymbol{R}$ に対して

$$T(\mathbf{x} + \mathbf{y}) = (a(x_1 + y_1)) = (ax_1) + (ay_1) = T(\mathbf{x}) + T(\mathbf{y}),$$
$$T(\alpha\mathbf{x}) = (\alpha a x_1) = \alpha(ax_1) = \alpha T(\mathbf{x})$$

より線形である．一方，後半の $T$ については，

$$\begin{aligned} T(\mathbf{x}) + T(\mathbf{y}) &= (ax_1 + b) + (ax_2 + b) = (a(x_1 + x_2) + 2b) \\ &\neq (a(x_1 + x_2) + b) = T(\mathbf{x} + \mathbf{y}) \end{aligned}$$

であるから，線形ではない．■

**例 1.3**　$R_\theta : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ を平面ベクトル $\mathbf{x} \in \boldsymbol{R}^2$ に，$\mathbf{x}$ を原点 $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ のまわりに角度 $\theta$ だけ回転させたベクトルを対応させる写像とする．これは，

$$R_\theta(\mathbf{x}) = \begin{pmatrix} x_1 \cos\theta - x_2 \sin\theta \\ x_1 \sin\theta + x_2 \cos\theta \end{pmatrix}$$

で与えられる．明らかに $R_\theta$ は線形である.

**問題 1.1**　例 1.3 を証明せよ.

**例 1.4**　$a_{jk} \in \boldsymbol{K}$, $j = 1, \cdots, M$, $k = 1, \cdots, N$ とする．$\mathbf{x} = (x_j) \in \boldsymbol{K}^N$ に対して

$$T(\mathbf{x}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1N}x_N \\ \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N \end{pmatrix} \tag{1.7}$$

とすると，$T$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像である．

**証明.** $\alpha \in \boldsymbol{K}$ に対して，

$$T(\alpha\mathbf{x}) = \begin{pmatrix} \alpha \sum_{n=1}^{N} a_{1n}x_n \\ \vdots \\ \alpha \sum_{n=1}^{N} a_{Mn}x_n \end{pmatrix} = \alpha T(\mathbf{x})$$

である．また，

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= \begin{pmatrix} \sum_{n=1}^{N} a_{1n}(x_n + y_n) \\ \vdots \\ \sum_{n=1}^{N} a_{Mn}(x_n + y_n) \end{pmatrix} \\ &= \begin{pmatrix} \sum_{n=1}^{N} a_{1n}x_n \\ \vdots \\ \sum_{n=1}^{N} a_{Mn}x_n \end{pmatrix} + \begin{pmatrix} \sum_{n=1}^{N} a_{1n}y_n \\ \vdots \\ \sum_{n=1}^{N} a_{Mn}y_n \end{pmatrix} \\ &= T(\mathbf{x}) + T(\mathbf{y}). \end{aligned}$$

■

後に証明されることであるが，$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像 $T$ はどのようなものでも必ずあるスカラー $a_{jk} \in \boldsymbol{K}$, $j = 1, \cdots, M$, $k = 1, \cdots, N$ により (1.7) の形に表される (第 3.1 節)．したがって線形写像は (1.7) の形のものに尽きているといえよう．しかし形がわかったとはいえ，議論が終わったわけではない．さら

に線形写像のもつ一般的な性質，あるいは $a_{jk}$ のとり方によって決まる線形写像の特性などを調べる必要がある．

本小節の最後に線形写像の中から有用なものを二つほど紹介しておく．

**例 1.5**（$N$ **点離散フーリエ変換**） 現在のディジタル技術を支える重要な線形写像の一つに $N$ 点離散フーリエ変換がある．これは次のように定義される．$\boldsymbol{K} = \boldsymbol{C}$ とする．$j, k \in \{1, 2, \cdots, N\}$ に対して

$$a_{jk} = \sqrt{\frac{1}{N}} \left\{ \cos\left(2\pi \frac{(k-1)(j-1)}{N}\right) - \sqrt{-1} \sin\left(2\pi \frac{(k-1)(j-1)}{N}\right) \right\} \tag{1.8}$$

とおく．(1.7) と同じ方法で

$$\mathcal{F}_N(\mathbf{x}) = \sqrt{N} \begin{pmatrix} a_{11}x_1 + \cdots + a_{1N}x_N \\ \vdots \\ a_{N1}x_1 + \cdots + a_{NN}x_N \end{pmatrix}$$

により $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像が定義される．この線形写像を $N$ **点離散フーリエ変換**あるいは単に**離散フーリエ変換**という．離散フーリエ変換の詳細は第 18 章で詳しく学ぶが，ここでは離散フーリエ変換のもつ意味を簡単に述べておく．(1.8) を用いて $N$ 個の $N$ 次元数ベクトル

$$\mathbf{S}_k = \begin{pmatrix} \overline{a_{1k}} \\ \vdots \\ \overline{a_{Nk}} \end{pmatrix} \in \boldsymbol{K}^N, \quad k = 1, \cdots, N$$

を定義する．任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して $\mathcal{F}_N(\mathbf{x}) \in \boldsymbol{K}^N$ を

$$\mathcal{F}_N(\mathbf{x}) = \begin{pmatrix} \mathcal{F}_N(\mathbf{x})_1 \\ \vdots \\ \mathcal{F}_N(\mathbf{x})_N \end{pmatrix}$$

により表す．このとき次のように $\mathbf{x}$ がベクトル $\mathbf{S}_1, \cdots, \mathbf{S}_N$ のスカラー倍の和で表されることが知られている．

$$\sqrt{N}\,\mathbf{x} = \mathcal{F}_N(\mathbf{x})_1\,\mathbf{S}_1 + \cdots + \mathcal{F}_N(\mathbf{x})_N\,\mathbf{S}_N.$$

(証明は第 18 章で行う)．この等式から $\mathcal{F}_N(\mathbf{x})_k$ は $\mathbf{x}$ にどのくらいの $\mathbf{S}_k$ が含まれているかを示す量になっていることがわかる．$\mathcal{F}_N(\mathbf{x})_1, \cdots, \mathcal{F}_N(\mathbf{x})_N$ は $\mathbf{x}$ の周

波数成分と呼ばれている．この変換の適用例を一つ挙げておく．図 1.5 の上図は図 1.1 であげた筆者の「アー」という音声の一部から作った離散信号 $\mathbf{x} \in \boldsymbol{K}^{410}$ である．この離散信号に 410 点離散フーリエ変換を施し，その値の絶対値 $|\mathcal{F}(\mathbf{x})_k|$，$k = 0, \cdots, 409$ をプロットしたものが図 1.5 の下図である．この図は音声信号 $\mathbf{x}$ がどのような周波数成分からなっているかを示している．

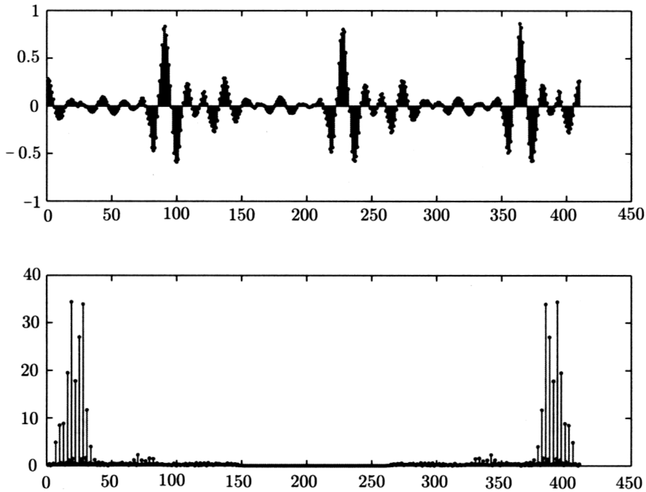



**図 1.5**　下図は上図のベクトルの離散フーリエ変換の絶対値のグラフ

**例 1.6**（$N$ **点離散コサイン変換**）　離散フーリエ変換と類似したものとして離散コサイン変換がある．$N$ を自然数とする．$m = 1, \cdots, N$ に対して

$$a_{jk} = c(k) \cos\left(\frac{\pi}{2N}(2k-1)(j-1)\right), \quad k = 1, \cdots, N$$

とおく．ただしここで

$$c(k) = \begin{cases} \sqrt{\dfrac{1}{N}} & (k = 1) \\ \sqrt{\dfrac{2}{N}} & (k = 2, \cdots, N) \end{cases}$$

である．このとき (1.7) により定義される線形写像

$$C_N(\mathbf{x}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1N}x_N \\ \vdots \\ a_{N1}x_1 + \cdots + a_{NN}x_N \end{pmatrix}$$

を，離散コサイン変換 (DCT, Discrete Cosine Transform) という．離散コサイン変換は画像圧縮 JPEG で実際に用いられている．

### 1.3.2 線形写像と連立 1 次方程式

線形写像と連立 1 次方程式との関連について述べておこう．次のような連立 1 次方程式を考える．

$$\begin{cases} a_{11}x_1 + \cdots + a_{1N}x_N = b_1 \\ \qquad\qquad \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N = b_M. \end{cases} \tag{1.9}$$

この連立 1 次方程式を解くとは，与えられた数 $b_1, \cdots, b_M$ に対して，(1.9) をみたす $x_1, \cdots, x_N$ を求めることである．このことは次のようにとらえることができる．$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への写像 $T$ を

$$T(\mathbf{x}) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1N}x_N \\ \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N \end{pmatrix}, \quad \mathbf{x} = (x_n)_{n=1,\cdots,N} \in \boldsymbol{K}^N$$

と定義する．例 1.4 より，$T$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像になっている．次の問題を考える．

> 出力データ $\mathbf{b}$ からどのような入力 $\mathbf{x} \in \boldsymbol{K}^N$ があったかを求めることはできないだろうか．

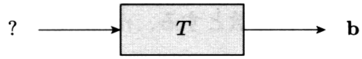



図 1.6

このような入力 $\mathbf{x}$ を出力 $\mathbf{b}$ から求めることが，すなわち連立 1 次方程式 (1.9) を解くことにほかならないのである．

さて，連立 1 次方程式は一般に解をもつ場合と解が存在しない場合とがある．さらに解をもつ場合でも，方程式によっては解がただ一つに限ることもあれば，解が複数存在することもある．線形代数では次のような問題が研究される．

- どのような場合，連立 1 次方程式 (1.9) は解をもつか？
- 解をもつ場合，解は一つに限るか，それとも複数存在するか？　複数存在する場合，すべての解はどのように記述されるか？

- 解が存在しない場合，$\mathbf{b}$ に最も近い値を出力できる入力データはどのようなものか？

本書ではこれらの問題も，これから先，扱うことになる．

# 第2章　行列と行列の演算

線形写像の構造を詳しく調べるために行列というものが使われる．本章では行列の定義をし，その例および基本的な性質を学ぶ．線形写像との関連については次章で解説する．

**ポイント2**　本章のポイントは行列の演算である．特に行列の積の定義は少し技巧的な規則で定義されているので注意してほしい．単位行列，対角行列，三角行列，転置，共役は後でしばしば使われるので，記憶に留めておくと便利である．

## 2.1　行列

まず行列の定義と行列の各部分の名称を述べておく．$MN$ 個の数 $a_{11}, \cdots, a_{MN}$ を

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} \tag{2.1}$$

のように並べたものを $(M, N)$ **行列**あるいは $M$ **行** $N$ **列の行列**という．特に $(N, N)$ 行列のことを $N$ **次正方行列**という．

行列 (2.1) において

$$a_{j1} \quad \cdots \quad a_{jN}$$

と並んだ部分をこの行列の第 $j$ 行，また

$$\begin{matrix} a_{1k} \\ \vdots \\ a_{Mk} \end{matrix}$$

と並んだ部分をこの行列の第 $k$ 列という．また $a_{jk}$ をこの行列の $(j,k)$ **成分**あるいは**第 $j$ 行，第 $k$ 列の成分**という．$(j,k)$ 成分を総称して単に成分という．便宜上 $a_{jk}$ を $a_{j,k}$ と表すこともある．たとえば $(j+m,\ j+n)$ 成分 $a_{j+m\,j+n}$ は見やすく表示するため $a_{j+m,\,j+n}$ と表すことがある．

記述の簡略化のため，行列 (2.1) をしばしば

$$(a_{jk})_{\substack{j=1,\cdots,M\\k=1,\cdots,N}} \quad \text{あるいは単に} \quad (a_{jk})_{j,k},\ (a_{jk})$$

と表す．

行列 $A=(a_{jk})$ の $(j,k)$ 成分を $A(j,k)$ と表すこともある．すなわち $A(j,k)=a_{jk}$ である．

各成分が $\boldsymbol{K}$ の要素からなる $(M,N)$ 行列全体のなす集合を $\boldsymbol{K}^{M\times N}$ と表す．集合の記号を用いて表すならば，

$$\boldsymbol{K}^{M\times N}=\left\{\begin{pmatrix} a_{11} & \cdots & a_{1N}\\ \vdots & \ddots & \vdots\\ a_{M1} & \cdots & a_{MN}\end{pmatrix} : a_{jk}\in\boldsymbol{K},\ j=1,\cdots,M,\ k=1,\cdots,N\right\}$$

である．$\boldsymbol{R}^{M\times N}$ に属する行列，すなわち各成分が実数からなる行列を**実行列**，また $\boldsymbol{C}^{M\times N}$ に属する行列，すなわち各成分が複素数である行列を**複素行列**という．

ところで $(M,1)$ 行列は $M$ 次元数ベクトルと同様の配列になっている．そこで $(M,1)$ 行列を $M$ 次元数ベクトルとみなす．また $(1,1)$ 行列 $(a_{11})$ は，しばしばスカラー $a_{11}$ と同一視する．すなわち

$$\boldsymbol{K}^{M\times 1}=\boldsymbol{K}^{M},$$
$$\boldsymbol{K}^{1\times 1}=\boldsymbol{K}^{1}=\boldsymbol{K}$$

とみなす．なお $(1,N)$ 行列を $N$ 次元横ベクトルとみなす．

本書では次の記法も用いる．$k,k+1,\cdots,k+l$ を $k:k+l$ と表す．行列 $A=(a_{mn})\in\boldsymbol{K}^{M\times N}$ に対して，

$$A\,(m:m+k,n:n+l)=\begin{pmatrix} a_{mn} & \cdots & a_{m,n+l}\\ \vdots & \ddots & \vdots\\ a_{m+k,n} & \cdots & a_{m+k,n+l}\end{pmatrix}\in\boldsymbol{K}^{(k+1)\times(l+1)}$$

と定める．また

$$A\,(m, n : n+k) = (a_{mn}, \cdots, a_{m,n+k}) \in \boldsymbol{K}^{1\times(k+1)},$$

$$A\,(m : m+k, n) = \begin{pmatrix} a_{mn} \\ \vdots \\ a_{m+k,n} \end{pmatrix} \in \boldsymbol{K}^{(k+1)\times 1} = \boldsymbol{K}^{k+1}$$

とし，特に

$$A\,(m, :) = A(m, 1:N) \in \boldsymbol{K}^{1\times N}, \quad A\,(:, n) = A\,(1:M, n) \in \boldsymbol{K}^{M\times 1}$$

とする．

このような記法はコロン記法といい，たとえばMATLAB® という技術計算ソフトのプログラム言語で使われている．

## 2.2 行列の演算

行列に関する解析を発展させるため，行列どうしの間に和，差，定数倍，積などの演算を定義する．このほかに転置，共役といった演算も定義する．これらの演算は行列の解析のための基礎である．

### 2.2.1 行列の和と差

二つの $(M, N)$ 行列 $A = (a_{jk})$, $B = (b_{jk})$ に対してその和，差をそれぞれ次のように定義する．

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1N}+b_{1N} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2N}+b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1}+b_{M1} & a_{M2}+a_{M2} & \cdots & a_{MN}+b_{MN} \end{pmatrix},$$

$$A - B = \begin{pmatrix} a_{11}-b_{11} & a_{12}-b_{12} & \cdots & a_{1N}-b_{1N} \\ a_{21}-b_{21} & a_{22}-b_{22} & \cdots & a_{2N}-b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1}-b_{M1} & a_{M2}-a_{M2} & \cdots & a_{MN}-b_{MN} \end{pmatrix}.$$

**例 2.1**

$$\begin{pmatrix} 1 & 2 & 3 \\ 6 & 4 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 8 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 3 \\ 8 & 7 & 10 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} 7 & 9 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} -6 & -6 \\ -3 & 3 \end{pmatrix}.$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} a+p \\ b+q \\ c+r \end{pmatrix}.$$

**注意 2.1**　二つの行列の和と差は，二つの行列の行の数と列の数がそれぞれ同じ場合にのみ定義されている．たとえば

$$\begin{pmatrix} 1 & 2 & 3 \\ 6 & 4 & 2 \\ 1 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

は定義されていない．

$(M,N)$ 行列 $A, B, C$ に対して，$A+B$ は $(M,N)$ 行列なので，この行列と $C$ との和 $(A+B)+C$ が定義できる．$A$ と $B+C$ の和 $A+(B+C)$ も定義できる．容易に

$$(A+B)+C = A+(B+C) \tag{2.2}$$

であることがわかる．(2.2) を

$$A+B+C$$

と表す．

この議論を繰り返して，$(M,N)$ 行列 $A_1, \cdots, A_p$ に対して，和

$$A_1 + \cdots + A_p$$

が定義される．$A_1 + \cdots + A_p$ を $\sum_{j=1}^{p} A_j$ とも表す．

### 2.2.2　行列の定数倍

$(M,N)$ 行列 $A = (a_{jk})$ とスカラー $\alpha \in \boldsymbol{K}$ に対して

$$\alpha A = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1N} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{M1} & \alpha a_{M2} & \cdots & \alpha a_{MN} \end{pmatrix}$$

と定める．$\alpha, \beta \in \boldsymbol{K}$ に対して，

$$(\alpha\beta)\, A = \alpha\,(\beta A)$$

であることが容易に確かめられる．

### 2.2.3　行列の積

行列 $A$ と $B$ の積 $AB$ の定義は少し技巧的である．行列の積 $AB$ は $A$ の列の数と $B$ の行の数が一致している場合に次のように定義される．

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{L1} & \cdots & a_{LM} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{M1} & \cdots & b_{MN} \end{pmatrix}$$

とするとき

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{L1} & a_{L2} & \cdots & a_{LM} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1N} \\ b_{21} & b_{22} & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{M1} & b_{M2} & \cdots & b_{MN} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{m=1}^{M} a_{1m}b_{m1} & \sum_{m=1}^{M} a_{1m}b_{m2} & \cdots & \sum_{m=1}^{M} a_{1m}b_{mN} \\ \sum_{m=1}^{M} a_{2m}b_{m1} & \sum_{m=1}^{M} a_{2m}b_{m2} & \cdots & \sum_{m=1}^{M} a_{2m}b_{mN} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{m=1}^{M} a_{Lm}b_{m1} & \sum_{m=1}^{M} a_{Lm}b_{m2} & \cdots & \sum_{m=1}^{M} a_{Lm}b_{mN} \end{pmatrix}.$$

これを $A$ と $B$ の**積**という．

たとえば，

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad B = \begin{pmatrix} x & z \\ y & w \end{pmatrix}$$

のときは,

$$AB = \begin{pmatrix} ax+cy & az+cw \\ bx+dy & bz+dw \end{pmatrix}$$

である.

一般に，定義より $(L,M)$ 行列 $A$ と $(M,N)$ 行列 $B$ の積 $AB$ は $(L,N)$ 行列であり，$AB$ の $(j,k)$ 成分は $\sum_{m=1}^{M} a_{jm}b_{mk}$ である.

なお，積 $AB$ では $A$ を $B$ に**左側から掛ける**，積 $BA$ では $A$ を $B$ に**右側から掛ける**という．一般に $AB$ と $BA$ は等しいとは限らない (問題 2.1 参照).

**注意 2.2**　$(N_1,N_2)$ 行列と $(N_3,N_4)$ 行列の積は $N_2=N_3$ の場合にのみ定義されている．たとえば

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

は定義されているが,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

は定義されていない.

**問題 2.1**　(1)　$A = (a \quad b)$, $B = \begin{pmatrix} c \\ d \end{pmatrix}$ とするとき，$AB$ および $BA$ を求めよ.

(2)　$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 2 & 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ のとき，$AB$, $BA$ を求めよ.

(3)　$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & -2 \end{pmatrix}$ のとき，$AB$, $BA$ を求めよ.

行列の演算に関するいくつかの代数的な性質を証明する.

**定理 2.1**　(1)　$A$ を $(N_1,N_2)$ 行列，$B$ を $(N_2,N_3)$ 行列，$C$ を $(N_3,N_4)$ 行列とする．このとき,

$$A(BC) = (AB)C$$

である．

(2)　$A$ を $(N_1, N_2)$ 行列，$B$, $C$ を $(N_2, N_3)$ 行列とすると

$$A(B+C) = AB + AC$$

が成り立つ．

(3)　$B$, $C$ を $(N_1, N_2)$ 行列，$A$ を $(N_2, N_3)$ 行列とすると

$$(B+C)A = BA + CA$$

が成り立つ．

**証明．** $A = (a_{\alpha\beta})$, $B = (b_{\gamma\delta})$, $C = (c_{\lambda\mu})$ とする．$BC = \left(\sum_{\upsilon=1}^{N_3} b_{\gamma\upsilon} c_{\upsilon\mu}\right)_{\gamma,\mu}$ である．したがって

$$A(BC) = \left(\sum_{\kappa=1}^{N_2} a_{\alpha\kappa} \left(\sum_{\upsilon=1}^{N_3} b_{\kappa\upsilon} c_{\upsilon\mu}\right)\right)_{\alpha,\mu}$$

$$= \left(\sum_{\upsilon=1}^{N_3} \left(\sum_{\kappa=1}^{N_2} a_{\alpha\kappa} b_{\kappa\upsilon}\right) c_{\upsilon\mu}\right)_{\alpha,\mu}.$$

一方，$AB = \left(\sum_{\kappa=1}^{N_2} a_{\alpha\kappa} b_{\kappa\delta}\right)_{\alpha,\delta}$ であるから，

$$(AB)\,C = \left(\sum_{\nu=1}^{N_3} \left(\sum_{\kappa=1}^{N_2} a_{\alpha\kappa} b_{\kappa\upsilon}\right) c_{\upsilon\mu}\right)_{\alpha,\mu}.$$

よって $A(BC) = (AB)C$ が示された．(2), (3) の証明は練習問題とする．■

$A$ が $(N_1, N_2)$ 行列, $B$ が $(N_2, N_3)$ 行列, $C$ が $(N_3, N_4)$ 行列であるとき $A$, $B$, $C$ の積を

$$ABC = (AB)C \ (= A(BC))$$

と定義する．

### 2.2.4　行列の積と線形写像

本小節では行列の積を用いて線形写像が作れることを示す．$A$ を $(M, N)$ 行列とする．$N$ 次元数ベクトル $\mathbf{x}$ は $(N, 1)$ 行列とみなせるので, $A$ と $\mathbf{x}$ の積 $A\mathbf{x} \in \boldsymbol{K}^{M\times 1}$ を定義できる．そこで $\mathbf{x}$ に $A\mathbf{x}$ を対応させる写像を定義し，これを $T_A$ により表す．すなわち

$$T_A(\mathbf{x}) = A\mathbf{x} \tag{2.3}$$

とする．明らかに $T_A$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への写像となっている．さらに次のことが成り立つ．

**定理 2.2**　$A \in \boldsymbol{K}^{M \times N}$ とする．(2.3) で定義される $T_A$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像である．

**証明.** $A = (a_{jk})$, $\mathbf{x} = (x_k)$ とすると

$$T_A(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{k=1}^{N} a_{1k} x_k \\ \vdots \\ \displaystyle\sum_{k=1}^{N} a_{Mk} x_k \end{pmatrix}$$

である．これは例 1.4 において定義された線形写像である．■

**定義 2.1**　$A \in \boldsymbol{K}^{M \times N}$ とする．(2.3) で定義される $T_A$ を $A$ **により定義される線形写像**あるいは $A$ **から定まる線形写像**という．

さて，行列 $A \in \boldsymbol{K}^{L \times M}$ と $B \in \boldsymbol{K}^{M \times N}$ に対して，これらの行列から定義される線形写像 $T_A, T_B$ はそれぞれ

$$T_A : \boldsymbol{K}^M \to \boldsymbol{K}^L, \quad T_B : \boldsymbol{K}^N \to \boldsymbol{K}^M$$

となっている．そこで $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$T_A \circ T_B(\mathbf{x}) = T_A(T_B(\mathbf{x}))$$

により $\boldsymbol{K}^N$ から $\boldsymbol{K}^L$ への写像を定義することができる．これを $T_A$ と $T_B$ の**合成写像**という．次のことが成り立つ．

**定理 2.3**　$A \in \boldsymbol{K}^{L \times M}$ と $B \in \boldsymbol{K}^{M \times N}$ とし，$C = AB \in \boldsymbol{K}^{L \times N}$ とする．このとき，

$$T_A \circ T_B = T_C$$

である．

**証明．** $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$T_A \circ T_B(\mathbf{x}) = T_A(T_B(\mathbf{x})) = T_A(B\mathbf{x}) = A(B\mathbf{x}) \\ = (AB)\mathbf{x} = T_C(\mathbf{x}).$$

■

このことから二つの行列 $A, B$ の積 $AB$ により定義される線形写像は，線形写像 $T_A$ と $T_B$ の合成写像 $T_A \circ T_B$ にほかならないことがわかる．

### 2.2.5 転置行列と共役行列

$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}$ に対して，

$$A^T = \begin{pmatrix} a_{11} & \cdots & a_{M1} \\ \vdots & \ddots & \vdots \\ a_{1N} & \cdots & a_{MN} \end{pmatrix}, \quad A^* = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{M1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1N}} & \cdots & \overline{a_{MN}} \end{pmatrix} \tag{2.4}$$

(ただし $\overline{a_{ij}}$ は $a_{ij}$ の複素共役) とする．$A^T$ を $A$ の転置，また $A^*$ を $A$ の共役という．$A^T$ のことを本によっては ${}^tA$ と表しているものも多い．$A$ が $(M,N)$ 行列ならば $A^T$, $A^*$ は $(N,M)$ 行列である．

$$\overline{A} = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{1N}} \\ \vdots & \ddots & \vdots \\ \overline{a_{M1}} & \cdots & \overline{a_{MN}} \end{pmatrix} \tag{2.5}$$

とし，これを $A$ の複素共役という．$A^* = \overline{A}^T$ である．明らかに

$$\left(A^T\right)^T = A, \quad (A^*)^* = A$$

である．転置，共役と行列の積の間には次の関係式が成り立つ．

**定理 2.4** $A$ を $(M,N)$ 行列，$B$ を $(N,L)$ 行列とする．このとき

$$(AB)^T = B^T A^T, \quad (AB)^* = B^* A^*.$$

**証明．** $A = (a_{ij})$, $A^T = (a'_{ij})$, $B = (b_{ij})$, $B^T = (b'_{ij})$ とする．$AB = (c_{ij})$ と

おくと

$$c_{rs} = \sum_{\mu=1}^{N} a_{r\mu} b_{\mu s}$$

である．一方，$(AB)^T = (c'_{ij})$ とすると

$$c'_{rs} = c_{sr} = \sum_{\mu=1}^{N} a_{s\mu} b_{\mu r} = \sum_{\mu=1}^{N} b'_{r\mu} a'_{\mu s}$$

であるから，$(AB)^T = B^T A^T$ が得られる．残りの主張も同様にして得られる．■

**問題 2.2**　$\mathbf{x} \in \boldsymbol{K}^N$ に対して，$\mathbf{x}^*\mathbf{x} \geq 0$ を示せ．$\mathbf{x}^*\mathbf{x} = 0$ ならば $\mathbf{x} = \mathbf{0}$ であることを示せ．

### 2.2.6　行列のアダマール積

二つの $(M, N)$ 行列

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{M1} & \cdots & b_{MN} \end{pmatrix}$$

に対して，

$$A_\bullet B = \begin{pmatrix} a_{11}b_{11} & \cdots & a_{1N}b_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1}b_{M1} & \cdots & a_{MN}b_{MN} \end{pmatrix}$$

とし，これを**アダマール積**という．

## 2.3　基本的な行列の例

行列は，その形によって特別な名称がつけられているものがある．そのようなもののうち本書でもしばしば現れる代表的な形を述べておきたい．

### 1．零行列

$(M, N)$ 行列で各成分が 0 になっているものを**零行列**といい，$O_{M\times N}$ で表す．すなわち

$$O_{M\times N} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

である．特に混乱の生じない場合は $O_{M\times N}$ を単に $O$ と表す．明らかに任意の $(M, L)$ 行列 $A$ に対して $O_{N\times M}A = O_{N\times L}$ であり，また $(L, M)$ 行列 $B$ に対して $BO_{M\times N} = O_{L\times N}$ である．

**2. 単位行列**

$(N, N)$ 行列で，$j = 1, \cdots, N$ に対して $(j, j)$ 成分が 1 で，$(j, k)$ 成分，$j \neq k$ が 0 であるような行列 $I_N$，すなわち

$$I_N = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

を $N$ **次単位行列**あるいは単に**単位行列**という．混乱のない場合 $I_N$ を単に $I$ と略記する．

$$\delta_{jk} = \begin{cases} 1 & (j = k) \\ 0 & (j \neq k) \end{cases} \tag{2.6}$$

と定義し，これを**クロネッカーの記号**あるいは**クロネッカーのデルタ記号**という．見やすくするため，$\delta_{jk}$ を $\delta_{j,k}$ と表すこともある．たとえば $\delta_{j+k\,m+n}$ の代わりに $\delta_{j+k,m+n}$ と表す．このクロネッカーのデルタ記号を用いれば，単位行列は

$$I_N = (\delta_{jk})$$

と表せる．

単位行列は次の性質をみたす．

**定理 2.5** $A$ を $(M, N)$ 行列とする．このとき

$$AI_N = A$$

が成り立つ．また $A$ が $(N, M)$ 行列のとき

$$I_N A = A$$

が成り立つ．特に $\mathbf{x} \in \boldsymbol{K}^N$ ならば

$$I_N \mathbf{x} = \mathbf{x}$$

である．

**証明.** $A = (a_{jk}) \in \boldsymbol{K}^{M\times N}$ とすると $AI_N$ の第 $(j,k)$ 成分は

$$\sum_{n=1}^{N} a_{jn}\delta_{nk} = a_{jk}\delta_{kk} = a_{jk}$$

である．ゆえに $AI_N = A$ が成り立つ．$A = (a_{jk}) \in \boldsymbol{K}^{N\times M}$ のとき $I_N A$ の第 $(j,k)$ 成分は

$$\sum_{n=1}^{N} \delta_{jn}a_{nk} = \delta_{jj}a_{jk} = a_{jk}$$

であるから $I_N A = A$ が得られる．$\mathbf{x} \in \boldsymbol{K}^N$ は $(N,1)$ 行列とみなせるので，$I_N\mathbf{x} = \mathbf{x}$ である．■

### 3. 対角行列

$(N,N)$ 行列で

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \lambda_N \end{pmatrix} \tag{2.7}$$

のように $j = 1, \cdots, N$ に対して $(j,j)$ 成分が $\lambda_j$ で，$(j,k)$ 成分，$j \neq k$ が $0$ であるような行列を $N$ **次対角行列**といい，

$$\mathrm{diag}\,(\lambda_1, \cdots, \lambda_N) \tag{2.8}$$

で表す．

正方行列でない行列に対しても対角行列に相当するものが次のように定義される．なお，今後二つの実数 $a, b$ に対して，$a \wedge b$ により $a \neq b$ ならば $a$ と $b$ の小さい方を表し，$a = b$ ならば $a$ を表すこととする．たとえば $3 \wedge 4 = 3$, $3 \wedge 3 = 3$ である．

$(M,N)$ 行列 $A = (a_{mn})$ で，$a_{jj}$, $1 \leq j \leq M \wedge N$ 以外の成分がすべて $0$ になっている行列を $(M,N)$ **対角行列**といい，

$$\mathrm{diag}_{M,N}(a_{11},\cdots,a_{M\wedge N,M\wedge N}) \tag{2.9}$$

と表す．たとえば

$$\mathrm{diag}_{2,3}(1,2)=\begin{pmatrix}1&0&0\\0&2&0\end{pmatrix},$$

$$\mathrm{diag}_{3,2}(1,2)=\begin{pmatrix}1&0\\0&2\\0&0\end{pmatrix}$$

はそれぞれ $(2,3)$ 対角行列，$(3,2)$ 対角行列の例である．

一般の $(M,N)$ 行列 $A=(a_{jk})$ に対して $a_{11},\cdots,a_{M\wedge N,\,M\wedge N}$ を $A$ の**対角成分**という．

## 4. 上三角行列と下三角行列

$(M,N)$ 行列 $A=(a_{jk})$ が

$$a_{jk}=0,\quad j>k$$

をみたすとき，$A$ を $(M,N)$ **上三角行列**あるいは単に**上三角行列**という．たとえば

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\0&a_{22}&a_{23}\\0&0&a_{33}\end{pmatrix},\quad\begin{pmatrix}a_{11}&a_{12}\\0&a_{22}\\0&0\end{pmatrix},\quad\begin{pmatrix}a_{11}&a_{12}&a_{13}\\0&a_{22}&a_{23}\end{pmatrix}$$

は上三角行列である．特に $(N,N)$ 行列でかつ上三角行列であるものを $N$ **次上三角行列**という．

一方，$(M,N)$ 行列 $A=(a_{jk})$ が

$$a_{jk}=0,\quad j<k$$

をみたすとき，$A$ を $(M,N)$ **下三角行列**あるいは単に**下三角行列**という．たとえば

$$\begin{pmatrix}a_{11}&0&0\\a_{21}&a_{22}&0\\a_{31}&a_{32}&a_{33}\end{pmatrix},\quad\begin{pmatrix}a_{11}&0\\a_{21}&a_{22}\\a_{31}&a_{32}\end{pmatrix},\quad\begin{pmatrix}a_{11}&0&0\\a_{21}&a_{22}&0\end{pmatrix}$$

は下三角行列である．特に $(N,N)$ 行列でかつ下三角行列であるものを $N$ **次下三角行列**という．

上三角行列あるいは下三角行列で，対角成分がすべて 1 の場合，すなわち $a_{kk}=1,\ k=1,\cdots,M\wedge N$ の場合，それぞれ**単位上三角行列**，**単位下三角行列**という．

本書では，$(M,N)$ 上三角行列で対角成分が $\lambda_1,\cdots,\lambda_{M\wedge N}$ であるものを

$$\mathrm{Trg}^{+}_{M,N}\,(\lambda_1,\cdots,\lambda_{M\wedge N})$$

と表し，$(M,N)$ 下三角行列で対角成分が $\lambda_1,\cdots,\lambda_{M\wedge N}$ であるものを

$$\mathrm{Trg}^{-}_{M,N}\,(\lambda_1,\cdots,\lambda_{M\wedge N})$$

と表す．

三角行列はたとえば連立方程式の解の研究に用いられる．

**問題 2.3**　(1)　$\begin{pmatrix} a & 0 \\ b & c \end{pmatrix}\begin{pmatrix} x & 0 \\ y & z \end{pmatrix}$ を求めよ．

(2)　$A=(a_{ij})$, $B=(b_{ij})$ を $N$ 次正方行列とする．$A$, $B$ が下三角行列のとき，$AB$ は下三角行列であることを示せ．$A$, $B$ が上三角行列のとき，$AB$ は上三角行列であることを示せ．またいずれの場合も $AB$ の対角成分は $a_{jj}b_{jj}$, $j=1,\cdots,N$ であることを示せ．

**5. その他**

本書を通して，$\mathbf{1}_{M\times N}$ により，すべての成分が 1 であるような $(M,N)$ 行列を表す．特に

$$\mathbf{1}_M=\mathbf{1}_{M\times 1}=\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}\in \boldsymbol{R}^M \tag{2.10}$$

とする．

## 2.4　逆行列

逆行列という考え方が線形代数では必要になるので，その定義と基本的な性質を述べておく．逆行列の定義をする．

**定義 2.2**　$N$ 次正方行列 $A$ に対して，ある $N$ 次正方行列 $B$ で

$$AB=I_N \quad \text{かつ} \quad BA=I_N \tag{2.11}$$

をみたすものが存在するとき，$B$ を $A$ の**逆行列**という．$A$ の逆行列が存在するとき，$A$ は**正則**あるいは**正則行列である**という．または $N$ **次正則行列**ともいう．

一般に逆行列は存在するとは限らない．たとえば零行列 $O_{N\times N}$ はどのような $B \in \boldsymbol{K}^{N\times N}$ をかけても $O_{N\times N}B = O_{N\times N}$ となる．したがって $O_{N\times N}$ の逆行列は存在しない．

また逆行列は，もし存在してもただ一つのみ存在する．このことを示しておこう．いま $A$ の逆行列 $B$ が存在するとする．$B'$ を $AB' = B'A = I_N$ をみたす $N$ 次正方行列とする．このとき

$$B = BI_N = B\left(AB'\right) = \left(BA\right)B' = I_N B' = B'$$

となる．

以下では，$N$ 次正則行列 $A$ に対して，$A$ の逆行列を $A^{-1}$ により表す．

**問題 2.4** $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$，$B = \begin{pmatrix} \frac{3}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$ とすると，

$$A^{-1} = B$$

であることを示せ．

逆行列については，本書では今後次のテーマを論ずることになる (第 4.4 節，第 6.2 節)．

**1)** **$A$ が正則であるかどうかの判定方法はないだろうか？**

**2)** **$A$ の逆行列を具体的に計算する方法あるいは公式はないだろうか？**

本節では逆行列の基本的な性質と，具体的な行列に対する逆行列の計算例をいくつか紹介する．

$\mathbf{x} \in \boldsymbol{K}^N$ を $\mathbf{x} \in \boldsymbol{K}^N$ に写す写像を**恒等写像** (identity) といい，$Id_N$ と表す．$N$ 次正則行列 $A$ から定義される線形写像 $T_A$ に対して，$T_{A^{-1}}$ は次の性質をもつ．

**定理 2.6** $A \in \boldsymbol{K}^{N\times N}$ を正則行列とする．このとき

$$T_A \circ T_{A^{-1}} = T_{A^{-1}} \circ T_A = Id_N.$$

**証明．** 定理 2.3 より

$$T_A \circ T_{A^{-1}} = T_{AA^{-1}} = T_{I_N},$$
$$T_{A^{-1}} \circ T_A = T_{A^{-1}A} = T_{I_N}$$

である．また定理 2.5 より，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$T_{I_N}(\mathbf{x}) = I_N \mathbf{x} = \mathbf{x}$$

であるから $T_{I_N} = Id_N$ である．■

**定理 2.7**　$N$ 次正方行列 $A$ に対して，ある $N$ 次正方行列 $B, C$ で，$AB = I_N, CA = I_N$ をみたすものが存在すれば，$A$ は正則行列で，$B = C = A^{-1}$ である．

**証明.** 仮定より

$$B = I_N B = (CA)B = C(AB) = CI_N = C$$

である．ゆえに $A$ は正則であり，$B = C = A^{-1}$ である．■

**注意 2.3**　後の章で証明されることであるが，じつは $N$ 次正方行列 $A$ に対して，ある $N$ 次正方行列 $B$ が存在し，$AB = I_N$ をみたすならば，$BA = I_N$ も成り立ち，したがって $B = A^{-1}$ となっている (系 6.4 参照).

**定理 2.8**　(1)　$A, B$ を $N$ 次正則行列とする．このとき，$AB$ も $N$ 次正則行列で，

$$(AB)^{-1} = B^{-1}A^{-1}$$

が成り立つ．

(2)　$A$ を $N$ 次正則行列とする．このとき $A^T, A^*$ も正則で，

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T, \quad (A^*)^{-1} = \left(A^{-1}\right)^*$$

が成り立つ．

**証明.** (1)　$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_N$ であり，また同様にして $(AB)B^{-1}A^{-1} = AA^{-1} = I_N$ である．ゆえに $AB$ は正則で，$(AB)^{-1} = B^{-1}A^{-1}$ である．

(2)　$I_N = I_N^T = (AA^{-1})^T = (A^{-1})^T A^T$ および $I_N = I_N^T = (A^{-1}A)^T = A^T (A^{-1})^T$ が成り立つ．ゆえに $\left(A^T\right)^{-1} = (A^{-1})^T$ が導かれる．同様にして $(A^*)^{-1} = (A^{-1})^*$ も得られる．■

最後に正則行列とその逆行列の例をいくつかあげておく．

**例 2.2** $N$ 次単位行列 $I_N$ は正則で，$I_N^{-1} = I_N$ である．

**証明.** $I_N I_N = I_N$ より明らかに $I_N^{-1} = I_N$ である． ■

**例 2.3** $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \boldsymbol{K}^{2\times 2}$ とする．$A$ が正則であるための必要十分条件は

$$ad - bc \neq 0$$

となることである．$ad - bc \neq 0$ のとき $A$ の逆行列は

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

である．

**証明.** $X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ として，$AX = I_2$ となるような $x, y, z, w$ を求めることを考えてみよう．

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{pmatrix}$$

であるから，

$$\begin{cases} ax + \ bz = 1 \ \cdots\cdots \text{(i)} \\ ay + bw = 0 \ \cdots\cdots \text{(ii)} \\ cx + \ dz = 0 \cdots\cdots \text{(iii)} \\ cy + dw = 1 \cdots\cdots \text{(iv)} \end{cases}$$

である．このとき

$$a \times \text{(iii)} - c \times \text{(i)} \text{ より} \qquad (ad - bc)\, z = -c \tag{2.12}$$

$$d \times \text{(i)} - b \times \text{(iii)} \text{ より} \qquad (ad - bc)\, x = d \tag{2.13}$$

$$a \times \text{(iv)} - c \times \text{(ii)} \text{ より} \qquad (ad - bc)\, w = a \tag{2.14}$$

$$d \times \text{(ii)} - b \times \text{(iv)} \text{ より} \qquad (ad - bc)\, y = -b \tag{2.15}$$

である．ここで (2.12) ~ (2.15) に現れる共通の因子 $ad - bc$ について次の二つの場合 (A), (B) が考えられる．

$$\text{(A)} : ad - bc = 0, \qquad \text{(B)} : ad - bc \neq 0.$$

(A) の場合，(2.12) ~ (2.15) から $c = d = a = b = 0$ となり，これは $AX = I_2$

に反する．すなわち $AX = I_2$ をみたす $X$ が存在するならば (B) が成り立つ．逆に (B) が成り立つ場合は，方程式 (2.12) ～ (2.15) を解けば

$$x = \frac{d}{ad-bc}, \qquad y = \frac{-b}{ad-bc},$$
$$z = \frac{-c}{ad-bc}, \qquad w = \frac{a}{ad-bc}$$

を得る．ここで

$$Y = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

とおくと，容易に $Y$ が $A$ の逆行列であること，すなわち

$$AY = YA = I_2$$

となることが確認できる．■

後の章で述べるが，行列の解析をするときの一つの常套手段として，行列を適切な行列との積をとって扱いやすい形の行列に変形するという方法がある．次にあげる例 2.4 および例 2.5 は行列を変形するときに用いられるものである．

**例 2.4**　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$, $\alpha \in \boldsymbol{K}$ に対して

$$E = I_N - \alpha\,\mathbf{x}\,\mathbf{y}^*$$

の形の $N$ 次正方行列を本書では**一般化基本行列**という．$\alpha\,\mathbf{y}^*\mathbf{x} \neq 1$ ならば $E$ は正則であり，

$$E^{-1} = I_N - \gamma\,\mathbf{x}\,\mathbf{y}^*, \quad \text{ただし} \quad \gamma = \frac{\alpha}{\alpha\,\mathbf{y}^*\mathbf{x} - 1}$$

である．したがって $E^{-1}$ も一般化基本行列である．

**証明.** $\beta = \alpha\,\mathbf{y}^*\mathbf{x}$ とおく．$\alpha^2\mathbf{x}\,\mathbf{y}^*\mathbf{x}\,\mathbf{y}^* = \alpha\,\mathbf{x}\,(\alpha\,\mathbf{y}^*\mathbf{x})\,\mathbf{y}^* = \alpha\beta\,\mathbf{x}\,\mathbf{y}^*$ であるから，

$$\begin{aligned}
&(I_N - \alpha\,\mathbf{x}\,\mathbf{y}^*)\left(I_N - \frac{\alpha}{\alpha\,\mathbf{y}^*\mathbf{x} - 1}\,\mathbf{x}\,\mathbf{y}^*\right) \\
&= I_N - \frac{\alpha}{\beta - 1}\mathbf{x}\,\mathbf{y}^* - \alpha\,\mathbf{x}\,\mathbf{y}^* + \frac{\alpha^2}{\beta - 1}\,\mathbf{x}\,\mathbf{y}^*\mathbf{x}\,\mathbf{y}^* \\
&= I_N - \frac{\alpha}{\beta - 1}\,\mathbf{x}\,\mathbf{y}^* - \alpha\,\mathbf{x}\,\mathbf{y}^* + \frac{\alpha\beta}{\beta - 1}\,\mathbf{x}\,\mathbf{y}^* = I_N.
\end{aligned}$$

また同様にして次のことも示せる．

$$\left(I_N - \frac{\alpha}{\alpha\,\mathbf{y}^*\mathbf{x} - 1}\,\mathbf{x}\,\mathbf{y}^*\right)(I_N - \alpha\,\mathbf{x}\,\mathbf{y}^*) = I_N.$$

■

**例 2.5**　$\mathbf{v} \in \boldsymbol{K}^N$, $\mathbf{v} \neq \mathbf{0}$ に対して

$$H_{\mathbf{v}} = I_N - \frac{2}{\mathbf{v}^*\mathbf{v}}\,\mathbf{v}\mathbf{v}^*$$

は一般化基本行列である．これを $\mathbf{v}$ **に関するハウスホルダー行列**あるいは $\mathbf{v}$ **に関するハウスホルダー変換**という．この行列は第 14.4 節で用いられる．

$$\frac{2}{\mathbf{v}^*\mathbf{v}}\,\mathbf{v}^*\mathbf{v} = 2 \neq 1$$

より $H_{\mathbf{v}}$ は正則である．また，$\alpha = \dfrac{2}{\mathbf{v}^*\mathbf{v}}$ とすると

$$\gamma = \frac{\alpha}{\alpha\mathbf{v}^*\mathbf{v} - 1} = \alpha$$

であるから，

$$H_{\mathbf{v}}^{-1} = H_{\mathbf{v}}$$

が成り立っている．さらに，$(\mathbf{v}\mathbf{v}^*)^* = (\mathbf{v}^*)^*\,\mathbf{v}^* = \mathbf{v}\mathbf{v}^*$ および $\alpha > 0$ より

$$\begin{aligned} H_{\mathbf{v}}^* &= (I_N - \alpha\,\mathbf{v}\mathbf{v}^*)^* = I_N^* - (\alpha\,\mathbf{v}\mathbf{v}^*)^* \\ &= I_N - \alpha\,(\mathbf{v}\mathbf{v}^*)^* = I_N - \alpha\,\mathbf{v}\mathbf{v}^*. \end{aligned}$$

ゆえに

$$H_{\mathbf{v}}^* = H_{\mathbf{v}}$$

も成り立っている．

# 第3章　線形写像と行列

本章では線形写像と行列の関係について述べる．行列から線形写像が構成できること，ならびに線形写像はすべて行列を用いて記述されることを示す．このことは線形代数の基本的な定理の一つである．

**ポイント3**　線形写像の表現行列，行列で定義される線形写像など，線形写像と行列の対応関係をしっかりと理解してほしい．また，線形写像の合成と表現行列の積との対応，線形写像の逆写像と表現行列の逆行列との対応なども重要である．

## 3.1　線形写像の行列による表現

定理 2.2 では行列の積を用いて線形写像がつくれることを示した．すなわち $A \in \boldsymbol{K}^{M \times N}$ に対して

$$T_A(\mathbf{x}) = A\mathbf{x}, \quad \mathbf{x} \in \boldsymbol{K}^N \tag{3.1}$$

により定義される線形写像が $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像になっている．じつはこの定理の逆が成り立っている．

**定理 3.1**　$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像とする．このときある $(M, N)$ 行列 $A \in \boldsymbol{K}^{M \times N}$ で

$$T(\mathbf{x}) = A\mathbf{x}, \quad x \in \boldsymbol{K}^N$$

をみたすものがただ一つとれる．すなわち $T = T_A$ となる $(M, N)$ 行列 $A$ が一意的に存在する．

**証明.** $\mathbf{x} \in \boldsymbol{K}^N$ は $N$ 次元基本ベクトル $\mathbf{e}_1^{(N)}, \cdots, \mathbf{e}_N^{(N)}$ により

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = x_1\mathbf{e}_1^{(N)} + \cdots + x_N\mathbf{e}_N^{(N)} = \sum_{n=1}^{N} x_n\mathbf{e}_n^{(N)}$$

と表せる．したがって，$T$ の線形性から

$$T(\mathbf{x}) = T\left(\sum_{n=1}^{N} x_n\mathbf{e}_n^{(N)}\right) = \sum_{n=1}^{N} x_n T(\mathbf{e}_n^{(N)}) \tag{3.2}$$

となっている．ここで各 $T(\mathbf{e}_n^{(N)})$ は $M$ 次元ベクトルであるから，それを

$$T(\mathbf{e}_n^{(N)}) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{Mn} \end{pmatrix}$$

とおくと，$M$ 次元基本ベクトル $\mathbf{e}_m^{(M)}, m = 1, \cdots, M$ により，各 $n$ に対して

$$T(\mathbf{e}_n^{(N)}) = a_{1n}\mathbf{e}_1^{(M)} + \cdots + a_{Mn}\mathbf{e}_M^{(M)} = \sum_{m=1}^{M} a_{mn}\mathbf{e}_m^{(M)}$$

と表せる．これを (3.2) に代入すると

$$\begin{aligned} T(\mathbf{x}) &= \sum_{n=1}^{N} x_n \left(\sum_{m=1}^{M} a_{mn}\mathbf{e}_m^{(M)}\right) = \sum_{m=1}^{M}\left(\sum_{n=1}^{N} a_{mn}x_n\right)\mathbf{e}_m^{(M)} \\ &= \begin{pmatrix} \displaystyle\sum_{n=1}^{N} a_{1n}x_n \\ \vdots \\ \displaystyle\sum_{n=1}^{N} a_{Mn}x_n \end{pmatrix} \end{aligned}$$

が得られる．一方，$A = (a_{ij}) \in \boldsymbol{K}^{M\times N}$ とすると

$$A\mathbf{x} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} \displaystyle\sum_{n=1}^{N} a_{1n}x_n \\ \vdots \\ \displaystyle\sum_{n=1}^{N} a_{Mn}x_n \end{pmatrix}$$

であるから，

$$T(\mathbf{x}) = A\mathbf{x}$$

となる．

次に，このような行列 $A$ がただ一つに限ることを証明しよう．もし $(M,N)$ 行列

$$A' = \begin{pmatrix} a'_{11} & \cdots & a'_{1N} \\ \vdots & \ddots & \vdots \\ a'_{M1} & \cdots & a'_{MN} \end{pmatrix}$$

により $T(\mathbf{x}) = A'\mathbf{x}$ と表せているとする．$n = 1, \cdots, N$ に対して

$$A\mathbf{e}_n^{(N)} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{Mn} \end{pmatrix}, \quad A'\mathbf{e}_n^{(N)} = \begin{pmatrix} a'_{1n} \\ \vdots \\ a'_{Mn} \end{pmatrix}$$

であるから，

$$\begin{pmatrix} a_{1n} - a'_{1n} \\ \vdots \\ a_{Mn} - a'_{Mn} \end{pmatrix} = A\mathbf{e}_n^{(N)} - A'\mathbf{e}_n^{(N)} = T\left(\mathbf{e}_n^{(N)}\right) - T\left(\mathbf{e}_n^{(N)}\right) = \mathbf{0}_M$$

である．よって $a_{mn} = a'_{mn}$ $(m = 1, \cdots, M)$ が得られる．■

**定義 3.1**　$(M,N)$ 行列 $A$ に対して，$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像 $T_A$ を $A$ によって**定義される線形写像**あるいは $A$ **から定まる線形写像**という．また，$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像 $T$ に対して定理 3.1 で得られる行列 $A$ を $T$ **の表現行列**，あるいは $T$ を**表現する行列**といい，$A_T$ と表す．

ここで定理 3.1 およびその証明が示していることをまとめておこう．

**(1)　線形写像がどのような構造を持っているかは，その表現行列の性質を調べればよい．**

**(2)　必要な性質をもつ線形写像を構成したいときには，行列を工夫して設計すればよい．**

**(3)　線形写像を表現する行列は次のように求めることが可能である．**

<u>ステップ 1.</u>　基本ベクトル $\mathbf{e}_k^{(N)}$ に対する出力

$$T(\mathbf{e}_k^{(N)}) = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{Mk} \end{pmatrix}, \quad k = 1, \cdots, N$$

を求める．

ステップ 2.

$$(a_{jk})_{j=1,\cdots,M,\ k=1,\cdots,N}$$

が求める行列 $A_T$ である．

**例題 3.1**　$\boldsymbol{K}^N \ni \mathbf{x}$ に対して

$$Id_N(\mathbf{x}) = \mathbf{x}$$

と定義される写像を**恒等写像**という．恒等写像は明らかに線形写像である．このとき，$Id_N$ を表現する行列は $I_N$ である．

**答.** $\mathbf{e}_k^{(N)} = Id_N(\mathbf{e}_k^{(N)})$ であるから，$Id_N$ を表現する行列は

$$\left(\mathbf{e}_1^{(N)}, \cdots, \mathbf{e}_N^{(N)}\right) = I_N$$

である．あるいは定理 2.6 と定理 3.1 の一意性からも証明することができる．■

**例題 3.2**　$a, b, c, d \in \boldsymbol{K}$ とする．

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}$$

により $\boldsymbol{K}^2$ から $\boldsymbol{K}^2$ への写像を定義する．$T$ が線形であることを示し，$T$ を表現する行列を求めよ．

**答.** $T$ が線形であることの証明は読者に委ねる．

$$T(\mathbf{e}_1^{(2)}) = T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}, \quad T(\mathbf{e}_2^{(2)}) = T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$$

であるから，$T$ を表す行列は $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ である．■

## 3.2　行列と線形写像の演算

線形写像どうしの和，スカラーとの積，合成などを定義し，それが行列の和，スカラーとの積，行列の積に対応していることを証明する．

いま $T_1, T_2$ を $\boldsymbol{K}^{N_2}$ から $\boldsymbol{K}^{N_1}$ への線形写像とするとき，$\mathbf{x} \in \boldsymbol{K}^{N_2}$ に対して

$$(T_1 + T_2)(\mathbf{x}) = T_1(\mathbf{x}) + T_2(\mathbf{x})$$

と定義する．明らかに $T_1 + T_2$ は $\boldsymbol{K}^{N_2}$ から $\boldsymbol{K}^{N_1}$ への線形写像である．$T_1, T_2$

を表現する $(N_1, N_2)$ 行列をそれぞれ $A_{T_1}$, $A_{T_2}$ とすると，

$$(T_1 + T_2)(\mathbf{x}) = A_{T_1}\mathbf{x} + A_{T_2}\mathbf{x} = (A_{T_1} + A_{T_2})\mathbf{x}$$

が成り立っている．したがって，$T_1 + T_2$ **を表現する行列は** $A_{T_1} + A_{T_2}$ **である．すなわち**

$$A_{T_1+T_2} = A_{T_1} + A_{T_2}.$$

$\alpha \in \boldsymbol{K}$ に対して，

$$(\alpha T_1)(\mathbf{x}) = \alpha T_1(\mathbf{x})$$

により，写像 $\alpha T_1$ を定義する．このとき，$\alpha T_1$ は線形であることが容易にわかる．また

$$(\alpha T_1)(\mathbf{x}) = \alpha T_1(\mathbf{x}) = \alpha A_{T_1}\mathbf{x}$$

であるから，$\alpha T_1$ **を表現する行列は** $\alpha A_{T_1}$ **である．すなわち**

$$A_{\alpha T_1} = \alpha A_{T_1}.$$

$S$ を $\boldsymbol{K}^{N_3}$ から $\boldsymbol{K}^{N_2}$ への線形写像，$T$ を $\boldsymbol{K}^{N_2}$ から $\boldsymbol{K}^{N_1}$ への線形写像とする．このとき，第 2.2.4 節ですでに述べたように合成写像 $T \circ S$ を

$$T \circ S(\mathbf{x}) = T(S(\mathbf{x})), \quad \mathbf{x} \in \boldsymbol{K}^{N_3}$$

により定義する．このとき $T \circ S$ は $\boldsymbol{K}^{N_3}$ から $\boldsymbol{K}^{N_1}$ への線形写像になっている．

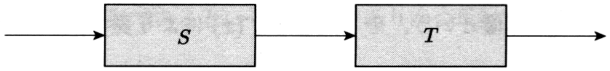



**図 3.1**

なぜならば

$$\begin{aligned} T \circ S(\alpha\mathbf{x}) &= T(\alpha S(\mathbf{x})) = \alpha T(S(\mathbf{x})) = \alpha T \circ S(\mathbf{x}) \\ T \circ S(\mathbf{x} + \mathbf{y}) &= T(S(\mathbf{x}) + S(\mathbf{y})) = T(S(\mathbf{x})) + T(S(\mathbf{y})) \\ &= T \circ S(\mathbf{x}) + T \circ S(\mathbf{y}) \end{aligned}$$

となっているからである．

定理 3.1 より，$S$ の表現行列 $A_S \in \boldsymbol{K}^{N_2 \times N_3}$, $T$ の表現行列 $A_T \in \boldsymbol{K}^{N_1 \times N_2}$, そして $T \circ S$ の表現行列 $A_{T \circ S} \in \boldsymbol{K}^{N_1 \times N_3}$ が存在する．定理 2.3 より

$$A_{T \circ S} = A_T A_S$$

であることがわかる．

$U$ を $\boldsymbol{K}^{N_3}$ から $\boldsymbol{K}^{N_2}$ への線形写像で，その表現行列を $A_U \in \boldsymbol{K}^{N_2 \times N_3}$ とすると

$$T \circ (S+U)(\mathbf{x}) = A_T(A_S + A_U)\mathbf{x}$$

である．したがって，$T \circ (S+U)$ を表現する行列は $A_T(A_S + A_U)$ である．

$R$ を $\boldsymbol{K}^{N_4}$ から $\boldsymbol{K}^{N_3}$ への線形写像とすると $A_R$ は $(N_3, N_4)$ 行列であり，

$$(S+U) \circ R(\mathbf{x}) = (A_S + A_U)A_R\mathbf{x}.$$

これより $(S+U) \circ R$ を表現する行列は $(A_S + A_U)A_R$ である．

**問題 3.1**

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}, \quad S\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a'x_1 + b'x_2 \\ c'x_1 + d'x_2 \end{pmatrix}$$

とするとき，$T \circ S$ の表現行列を求めよ．

## 3.3　写像と逆写像

はじめに写像と逆写像について基本事項を説明する．われわれが扱うのは線形写像とその逆写像であるが，ここでは一般の場合について述べておきたい．

一般に，集合 $X$ の要素 $x$ に対して，ある規則で集合 $Y$ の要素 $y$ を定める対応 $T$ を $X$ から $Y$ への写像という．そして $y$ を $T(x)$ により表す．$T$ が $X$ から $Y$ への写像であることを

$$T : X \to Y$$

なる記号で表す．また $T$ が $x \in X$ にある規則で $y \in Y$ を対応させる写像であることを

$$T : X \ni x \mapsto y \in Y$$

と表す．たとえば $A \in \boldsymbol{K}^{M \times N}$ のとき，$A$ から定まる線形写像 $T_A$ は

$$T_A : \boldsymbol{K}^N \ni \mathbf{x} \mapsto A\mathbf{x} \in \boldsymbol{K}^M$$

のように表される．

集合 $X$ から集合 $X$ への写像で $X$ の要素 $x$ に $x$ を対応させる写像を $X$ 上の恒等写像といい，$Id_X$ と表す．集合 $X$ から集合 $Y$ への写像 $T$ が**逆写像をもつ**とは，

$Y$ から $X$ への写像 $S$ で,

$$T \circ S = Id_Y, \quad S \circ T = Id_X$$

をみたすものが存在することである．このような写像 $S$ を $T$ の**逆写像**という．

逆写像は，一般に存在するとは限らないが，もし存在する場合はただ一つしか存在しない．なぜならば，$Y$ から $X$ への写像 $S'$ が $T \circ S' = Id_Y$, $S' \circ T = Id_X$ をみたしているならば,

$$\begin{aligned} S' &= S' \circ Id_Y = S' \circ (T \circ S) = (S' \circ T) \circ S \\ &= Id_X \circ S = S \end{aligned}$$

となっているからである．そこで以下では $T$ の逆写像が存在するとき，$T$ の逆写像を $T^{-1}$ により表す．

逆写像が存在するための有用な必要十分条件を述べておく．そのため，まず全射，単射，全単射の定義をする．

**定義 3.2**　集合 $X$ から集合 $Y$ への写像 $T$ に対して,

$$\mathcal{R}(T) = \{T\,(x) : x \in X\}$$

を $T$ の**値域**あるいは**像空間**という．$\mathcal{R}(T) = Y$ が成り立つとき，$T$ は**全射**であるという．また

$$T(x) = T(x') \quad \text{ならば} \quad x = x'$$

をみたすとき, $T$ は**単射**であるという．全射でありかつ単射である写像を**全単射**という．

次のことが成り立つ.

**命題 3.2**　$T$ を集合 $X$ から集合 $Y$ への写像とする．$T$ が逆写像をもつための必要十分条件は，$T$ が全単射となることである．

**証明.** $T$ が全単射であるとする．全射であることより，任意の $z \in Y$ に対して，$T(x) = z$ となる $x \in X$ が存在するが，さらに単射であることより，このような $x$ はただ一つしか存在しないことがわかる．実際 $T(x) = z = T(x')$ ならば $x = x'$ である．そこで $z \in Y$ に対して $T(x) = z$ をみたす $x$ を $S(z)$ で表し，$z$ に対して $S(z)$ を対応させる写像を $S$ とする．このとき

$$T(S(z)) = T(x) = z,$$
$$S(T(x)) = S(z) = x$$

をみたす．ゆえに $S$ は $T$ の逆写像である．

一方，$T$ の逆写像 $S$ が存在するとき，任意の $z \in Y$ に対して，

$$z = T(S(z))$$

であるから，$T$ は全射である．また $T(x) = T(x')$ であれば

$$x = S(T(x)) = S(T(x')) = x'$$

が成り立ち，$T$ が単射であることもわかる．■

線形写像の場合，単射であるための次の必要十分条件は有用である．

**命題 3.3** $T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像とし，

$$\mathcal{N}(T) = \left\{\mathbf{x} : \mathbf{x} \in \boldsymbol{K}^N,\ T(\mathbf{x}) = \mathbf{0}_M\right\}$$

とする．$T$ が単射であるための必要十分条件は $\mathcal{N}(T) = \{\mathbf{0}_N\}$ となることである．

**証明．** 命題 1.1 より $\mathbf{0}_N \in \mathcal{N}(T)$ である．$T$ が単射であるとする．このとき，$\mathbf{x} \in \mathcal{N}(T)$ ならば，$T(\mathbf{x}) = \mathbf{0}_M = T(\mathbf{0}_N)$ であるから $\mathbf{x} = \mathbf{0}_N$ である．ゆえに $\mathcal{N}(T) = \{\mathbf{0}_N\}$ が成り立つ．逆に $\mathcal{N}(T) = \{\mathbf{0}_N\}$ であるとする．もし $T(\mathbf{x}) = T(\mathbf{x}')$ であるとすると，$T$ が線形であることより，

$$T(\mathbf{x} - \mathbf{x}') = T(\mathbf{x}) - T(\mathbf{x}') = \mathbf{0}_M$$

であるから，$\mathbf{x} - \mathbf{x}' = \mathbf{0}_N$ が成り立つ．すなわち $T$ は単射である．■

**定義 3.3** 命題 3.3 で定義された $\mathcal{N}(T)$ を $T$ の**核**という．

本書を通して次の記号を用いる．

**定義 3.4** $(M, N)$ 行列 $A$ に対して，

$$\mathcal{R}(A) = \mathcal{R}(T_A),$$
$$\mathcal{N}(A) = \mathcal{N}(T_A)$$

と定義する．$\mathcal{R}(A)$ を $A$ の**値域**あるいは**像空間**，また $\mathcal{N}(A)$ を $A$ の**核**という．

さて，線形写像の逆写像は次の性質をもつ．

**定理 3.4**　$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像で，逆写像をもつとする．このとき $T$ の逆写像 $T^{-1}$ は $\boldsymbol{K}^M$ から $\boldsymbol{K}^N$ への線形写像である．

**証明.** $\mathbf{z} \in \boldsymbol{K}^M$ に対して，$T^{-1}(\mathbf{z}) = \mathbf{x}$ とすると，$T(\mathbf{x}) = \mathbf{z}$ であり，また $\alpha \in \boldsymbol{K}$ に対して $\alpha T^{-1}(\mathbf{z}) = \alpha \mathbf{x}$ であるから，

$$T\left(\alpha T^{-1}(\mathbf{z})\right) = T(\alpha \mathbf{x}) = \alpha T(\mathbf{x}) = \alpha \mathbf{z} = T\left(T^{-1}(\alpha \mathbf{z})\right)$$

が成り立つ．$T$ が単射であることより，$\alpha T^{-1}(\mathbf{z}) = T^{-1}(\alpha \mathbf{z})$ を得る．$\mathbf{z}, \mathbf{w} \in \boldsymbol{K}^M$ に対して，$T^{-1}(\mathbf{z}) = \mathbf{x}, T^{-1}(\mathbf{w}) = \mathbf{y}$ とおく．このとき $T(\mathbf{x}) = \mathbf{z}, T(\mathbf{y}) = \mathbf{w}$ より

$$\begin{aligned} T\left(T^{-1}(\mathbf{z}) + T^{-1}(\mathbf{w})\right) &= T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \\ &= \mathbf{z} + \mathbf{w} = T\left(T^{-1}(\mathbf{z}+\mathbf{w})\right). \end{aligned}$$

ここで，$T$ が単射であることより

$$T^{-1}(\mathbf{z}) + T^{-1}(\mathbf{w}) = T^{-1}(\mathbf{z}+\mathbf{w})$$

が成り立つ．■

定理 3.4 から線形写像 $T$ が逆写像をもつとき，$T^{-1}$ を表現する行列が存在することがわかる．次の定理が成り立つ．

**定理 3.5**　$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像とし，$T$ を表現する行列を $A$ とする．$T$ が逆写像をもつための必要十分条件は，$A$ が正則となることである．さらにこのとき，

$$T^{-1} = T_{A^{-1}}$$

である．

**証明.** $T$ が逆写像 $T^{-1}$ をもつとする．$T$ を表現する行列を $A$，$T^{-1}$ を表現する行列を $B$ とすると，

$$AB\mathbf{x} = T\left(T^{-1}(\mathbf{x})\right) = \mathbf{x} = T^{-1}(T(\mathbf{x})) = BA\mathbf{x}$$

である．したがって $AB = I_N = BA$ であるから $B = A^{-1}$ である．

十分性を示す．任意の $\mathbf{z} \in \boldsymbol{K}^N$ に対して，$\mathbf{x} = A^{-1}\mathbf{z}$ とおくと，

$$T(\mathbf{x}) = A\mathbf{x} = AA^{-1}\mathbf{z} = \mathbf{z}$$

となっている．ゆえに $T$ は全射である．また $T(\mathbf{x}) = T(\mathbf{y})$ とすると，

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}T(\mathbf{x}) = A^{-1}T(\mathbf{y}) = A^{-1}A\mathbf{y} = \mathbf{y}$$

より $T$ は単射になっている．これで十分性が示せた．以上の証明から $T^{-1} = T_{A^{-1}}$ であることもわかる． ■

# 第4章　ガウスの消去法

本章では連立1次方程式の解を求めるための強力な方法であるガウスの消去法について述べる．また，ガウスの消去法を用いて具体的に逆行列の計算を行う方法も述べる．

まず連立1次方程式

$$\left\{\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= b_1 \\ \vdots \qquad\qquad &\quad \vdots \\ a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N &= b_M \end{aligned}\right. \tag{4.1}$$

に関する基本的な用語を定めておく．与えられたスカラー $a_{11}, \cdots, a_{MN}$ および $b_1, \cdots, b_M$ に対して (4.1) をみたす $x_1, \cdots, x_N$ を求めることを**連立1次方程式(4.1) を解く**という．また (4.1) をみたす $x_1, \cdots, x_N$ を (4.1) の**解**という．連立1次方程式 (4.1) の係数から作った行列

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} \tag{4.2}$$

を連立1次方程式 (4.1) の**係数行列**という．また

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1N} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{M1} & \cdots & a_{MN} & b_M \end{array}\right) \tag{4.3}$$

のように配列したものを**拡大係数行列**という．

なお連立方程式 (4.1) は $A$ を係数行列 (4.2) とし，

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}$$

とすると，簡単に

$$A\mathbf{x} = \mathbf{b} \tag{4.4}$$

と表すことができる．

本章では**方程式の数と未知数の数が同じであり，しかもただ一つの解をもつ場合**についてガウスの消去法による解法を述べる．解が複数存在する場合や，方程式の数と未知数の数が異なる場合については後の章で論ずる．

**ポイント 4**　具体的な連立 1 次方程式の解や行列の逆行列が計算できるようにしてほしい．実際に問題を解いて慣れておくとよいだろう．

## 4.1　具体例

はじめに具体例を用いてガウスの消去法とはどのようなものかを述べる．

**例 4.1**　連立 1 次方程式

$$\begin{cases} x_1 + x_2 - x_3 = 1 \cdots\cdots (1) \\ 2x_1 + x_2 - x_3 = 0 \cdots\cdots (2) \\ 2x_1 + 4x_2 - 3x_3 = 1 \cdots\cdots (3) \end{cases} \tag{4.5}$$

を解くことを考える．

まず第 1 式の 2 倍を第 2 式から引き (このことを $-2 \times (1) + (2)$ と表す)，次に第 1 式の 2 倍を第 3 式から引く．言い換えれば $-2 \times (1) + (3)$ を計算する．これより第 2 式と第 3 式から $x_1$ が消去され，次式が得られる．

$$\begin{cases} x_1 + x_2 - x_3 = 1 \cdots\cdots (1) \\ \quad - x_2 + x_3 = -2 \cdots\cdots (4) \\ \quad 2x_2 - x_3 = -1 \cdots\cdots (5) \end{cases}$$

次に $2 \times (4) + (5)$ を計算して，第 3 式から $x_2$ を消去する．その式を (6) とすると，次のようになる．

$$\begin{cases} x_1 + x_2 - x_3 = \ 1 \ \cdots\cdots \ (1) \\ \quad\ - x_2 + x_3 = -2 \cdots\cdots \ (4) \\ \qquad\qquad\ x_3 = -5 \cdots\cdots \ (6) \end{cases} \tag{4.6}$$

これにより結局，連立方程式 (4.5) は係数行列が上三角行列の連立方程式 (4.6) に変形された．このように，ある行を定数倍して他の行に加えるなどして係数行列を上三角行列に変形する方法を**ガウスの消去法**という (正確には，後に定義する行列の行に関する基本変形を用いて，行列を上三角行列に変形することである)．

さて，このように変形しておけば，解は次のような操作で容易に求めることができる．まず (6) より

$$x_3 = -5$$

が得られる．次に (4) にこれを代入して，

$$x_2 = -(-2 - x_3) = -3$$

が得られる．これらを (1) に代入すれば

$$x_1 = 1 - x_2 + x_3 = -1$$

となる．したがって $x_1, x_2, x_3$ が (4.5) をみたすならば，

$$x_1 = -1, \quad x_2 = -3, \quad x_3 = -5$$

であることがわかる．

逆に，これが解になっていることは容易に確かめることができる．

このようにして解を求める方法を**後退代入法**という．

さて，以上の操作を連立 1 次方程式の係数のみを用いた略記法で記してみよう．この略記法を知っておくと，連立 1 次方程式を解く際に非常に便利である．まず連立 1 次方程式 (4.5) の係数行列を

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 4 & -3 \end{pmatrix}$$

とする．次に，与えられたデータ $1, 0, 1$ も加えて拡大係数行列

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & 1 & -1 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right)$$

をつくる．ガウスの消去法の最初のステップは，(1) 式を $-2$ 倍して (2), (3) に加えることにより (4), (5) を作るというもので，この部分を次のように表す．

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & 1 & -1 & 0 \\ 2 & 4 & -3 & 1 \end{array}\right) \xrightarrow[-2R_1+R_2,\ -2R_1+R_3]{} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & 2 & -1 & -1 \end{array}\right).$$

ここで $-2R_1 + R_2$ は矢印の左にある表の第 1 行の $-2$ 倍を第 2 行に加えることを意味する．また，$-2R_1 + R_3$ は矢印の左にある表の第 1 行の $-2$ 倍を第 3 行に加えることを意味する．次にガウスの消去法では (4) を 2 倍して (5) に加えて (6) を作った．この部分は，

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & 2 & -1 & -1 \end{array}\right) \xrightarrow[2R_2+R_3]{} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & 1 & -5 \end{array}\right)$$

と表せる．ただし，ここで $2R_2 + R_3$ は矢印の左にある表の第 2 行の 2 倍を第 3 行に加えることを意味している．

以上の結果，係数行列は上三角行列

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

となるようにできた．そして上三角行列が得られると，方程式の下側の式を使って変数を消去し，解を逐次求めていくことになる．

ガウスの消去法を用いて連立 1 次方程式を解く例をもう一つ述べておこう．

**例 4.2**

$$\begin{cases} \phantom{2x_1 - {}} x_2 - \phantom{2}x_3 = 1 \cdots\cdots & (1) \\ 2x_1 - 2x_2 + 2x_3 = 0 \cdots\cdots & (2) \\ \phantom{2}x_1 + 2x_2 - \phantom{2}x_3 = 1 \cdots\cdots & (3) \end{cases}$$

まず (1) と (2) を入れ替えて次のようにする．

$$\begin{cases} 2x_1 - 2x_2 + 2x_3 = 0 \cdots\cdots & (4) \\ \phantom{2x_1 - {}} x_2 - \phantom{2}x_3 = 1 \cdots\cdots & (5) \\ \phantom{2}x_1 + 2x_2 - \phantom{2}x_3 = 1 \cdots\cdots & (3) \end{cases}$$

(4) を $\dfrac{1}{2}$ 倍して

$$\begin{cases} x_1 - \phantom{2}x_2 + x_3 = 0 \cdots\cdots & (6) \\ \phantom{x_1 + 2}x_2 - x_3 = 1 \cdots\cdots & (5) \\ x_1 + 2x_2 - x_3 = 1 \cdots\cdots & (3) \end{cases}$$

ここで $-(6)+(3)$ とすると

$$\begin{cases} x_1 - \phantom{3}x_2 + \phantom{2}x_3 = 0 \cdots\cdots & (6) \\ \phantom{x_1 - 3}x_2 - \phantom{2}x_3 = 1 \cdots\cdots & (5) \\ \phantom{x_1 - }3x_2 - 2x_3 = 1 \cdots\cdots & (7) \end{cases}$$

次に $(7) - 3 \times (5)$ を計算して

$$\begin{cases} x_1 - x_2 + x_3 = \phantom{-}0 \ \cdots\cdots & (6) \\ \phantom{x_1 - }x_2 - x_3 = \phantom{-}1 \ \cdots\cdots & (5) \\ \phantom{x_1 - x_2 - }x_3 = -2 \cdots\cdots & (8) \end{cases}$$

この変形をすでに述べた略記法で書けば次のようになる．ただし $\underset{\alpha R_j}{\longrightarrow}$ は矢印の左側にある表の第 $j$ 行を $\alpha$ 倍する操作を，そして $R_j \longleftrightarrow R_k$ は矢印の左の行列の第 $j$ 行と第 $k$ 行を交換する操作を表す．

$$\left(\begin{array}{ccc|c} 0 & 1 & -1 & 1 \\ 2 & -2 & 2 & 0 \\ 1 & 2 & -1 & 1 \end{array}\right) \underset{R_1 \leftrightarrow R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 2 & -2 & 2 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 2 & -1 & 1 \end{array}\right) \underset{\frac{1}{2}R_1}{\longrightarrow}$$

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 2 & -1 & 1 \end{array}\right) \underset{-R_1+R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 3 & -2 & 1 \end{array}\right)$$

$$\underset{-3R_2+R_3}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right).$$

これより，

$$x_3 = -2$$

が得られる．すると (5) にこれを代入して

$$x_2 = 1 + x_3 = -1.$$

さらにこれらを (6) に代入して

$$x_1 = x_2 - x_3 = 1.$$

したがって求める解は

$$x_1 = 1, \quad x_2 = -1, \quad x_3 = -2$$

である．

例 4.1 と例 4.2 はともにガウスの消去法に基づく解法である．ただしこの二つの例には違いがある．それは例 4.2 では一回目の操作において第 1 式の定数倍を第 2, 3 式に加えていないことである．例 4.2 の場合，第 1 式の $x_1$ の係数が 0 であるため，この式を定数倍して加えても第 2, 3 式の $x_1$ の項を消去できないからである．そこで，第 1 式と $x_1$ の係数が 0 でない式と入れ替えることから始めている．この操作を**ピボット選択**という．ピボットについての詳しい解説は後述する (第 11.2 節).

**式変形の略記法のまとめ**

一般に

$$A \underset{\alpha R_j + R_k}{\longrightarrow} B$$

は矢印の左側にある表 $A$ の第 $j$ 行の $\alpha$ 倍に第 $k$ 行をたすと矢印の右側の表 $B$ になること意味している ($R$ は row(行) の頭文字)．また

$$A \underset{\alpha R_j + R_k,\ \beta R_l + R_m}{\longrightarrow} B$$

は矢印の左側の表の第 $j$ 行の $\alpha$ 倍を第 $k$ 行に加え，その後第 $l$ 行の $\beta$ 倍を第 $m$ 行に加えると矢印の右側の表 $B$ になることを意味している．

$$A \underset{R_j \longleftrightarrow R_k}{\longrightarrow} B$$

は $A$ の第 $j$ 行と第 $k$ 行を交換することを意味している．

## 4.2 より一般の場合のガウスの消去法

一般の場合のガウスの消去法について述べる．基本的な方針は前節のものと同様で，行に関する変形をして，係数行列が上三角行列の形になるようにして，それから後退代入法により解を求めるというものである．まず第一段階として上三角行列への変形から始める．

### 4.2.1 上三角行列への変形

与えられた $\mathbf{y} \in \boldsymbol{K}^N$ に対して

$$
\begin{cases}
a_{11}x_1 + \cdots + a_{1N}x_N = y_1 & \cdots (1) \\
\qquad\qquad \vdots & \quad \vdots \\
a_{N1}x_1 + \cdots + a_{NN}x_N = y_N & \cdots (N)
\end{cases}
$$

をみたす $\mathbf{x} \in \boldsymbol{K}^N$ を求めることを考える．ここでも記述を簡単にするため，係数のみを書いた表

$$
\left(\begin{array}{ccc|c}
a_{11} & \cdots & a_{1N} & y_1 \\
\vdots & \ddots & \vdots & \vdots \\
a_{N1} & \cdots & a_{NN} & y_N
\end{array}\right)
$$

を用いる．

まず $a_{11}, \cdots a_{N1}$ の中で 0 でないものがあるかどうかを調べる．もしすべて 0 ならば操作を停止する．この場合の解の求め方については第 12 章で学ぶ．ここでは $a_{11}, \cdots a_{N1}$ の中で 0 でないものがある場合を扱う．0 でないものの一つ $a_{j1}$ を選び，第 $j$ 行と第 1 行と入れ替える ($j = 1$ の場合は入れ替えの操作はしない．なお計算機による数値計算を行う場合は，桁落ちを避けるため $|a_{j1}|$ の中で最大値をとるものを選ぶようにする)．その結果を

$$
\left(\begin{array}{ccc|c}
a_{11}^{(1)} & \cdots & a_{1N}^{(1)} & y_1^{(1)} \\
\vdots & \ddots & \vdots & \vdots \\
a_{N1}^{(1)} & \cdots & a_{NN}^{(1)} & y_N^{(1)}
\end{array}\right)
$$

と表す．ここで ($a_{i1}^{(1)} \neq 0$ なる第 $i$ 行に対して)，

$$
(\text{第 } i \text{ 行}) - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \times (\text{第 } 1 \text{ 行})
$$

を計算すると

$$
\left(\begin{array}{cccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1N}^{(1)} & y_1^{(1)} \\
0 & a_{22}^{(1)\prime} & \cdots & a_{2N}^{(1)\prime} & y_2^{(1)\prime} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & a_{N2}^{(1)\prime} & \cdots & a_{NN}^{(1)\prime} & y_N^{(1)\prime}
\end{array}\right)
$$

の形の結果が得られる．

次に $a_{i2}^{(1)\prime}$ $(2 \leq i \leq N)$ の中で 0 にならないものがあるかどうか調べる．もしす

べて0なら操作を停止する (すべて0になるような場合は第12章で扱う). ここでは $a_{i2}^{(1)\prime}$ $(2 \leq i \leq N)$ の中に0でないものがある場合を扱う. 0でないものの一つ $a_{j2}^{(1)\prime}$ を選び, 第2行と第 $j$ 行を入れ替える (たとえば数値計算のためには $\left|a_{j2}^{(1)\prime}\right|$ の中で最大値をとるものを選ぶ). その結果を

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1N}^{(1)} & y_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2N}^{(2)} & y_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{N2}^{(2)} & \cdots & a_{NN}^{(2)} & y_N^{(2)} \end{array}\right)$$

とする. そして $a_{i2}^{(2)} \neq 0$ なる第 $i$ 行 $(3 \leq i \leq N)$ に対して,

$$(\text{第}\ i\ \text{行}) - \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \times (\text{第}\ 2\ \text{行})$$

を計算すると

$$\left(\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1N}^{(1)} & y_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2N}^{(2)} & y_2^{(2)} \\ 0 & 0 & a_{33}^{(2)\prime} & \cdots & a_{3N}^{(2)\prime} & y_3^{(2)\prime} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{N3}^{(2)\prime} & \cdots & a_{NN}^{(2)\prime} & y_N^{(2)\prime} \end{array}\right)$$

を得る.

この操作を繰り返して, 途中で停止することなく最後まで実行できる場合,

$$\left(\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & a_{1N}^{(1)} & y_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & \cdots & a_{2N}^{(2)} & y_2^{(2)} \\ \vdots & 0 & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & a_{N-1,N-1}^{(N-1)} & a_{N-1,N}^{(N-1)} & y_{N-1}^{(N-1)} \\ 0 & 0 & \cdots & 0 & a_{NN}^{(N)} & y_N^{(N)} \end{array}\right)$$

$$\text{ただし,}\quad a_{11}^{(1)} a_{22}^{(2)} \cdots a_{NN}^{(N)} \neq 0$$

の形の結果となる. これにより問題は上三角行列の場合に帰着できた.

次に, 上三角行列が得られた後の解法を述べる.

### 4.2.2 前進代入法と後退代入法

一般に与えられた $\mathbf{y} \in \boldsymbol{K}^N$ に対して

$$
\begin{pmatrix}
a_{11} & a_{12} & \cdots & \cdots & a_{1N} \\
0 & a_{22} & \cdots & \cdots & a_{2N} \\
\vdots & 0 & \ddots & & \vdots \\
\vdots & \vdots & \ddots & a_{N-1,N-1} & a_{N-1,N} \\
0 & 0 & \cdots & 0 & a_{NN}
\end{pmatrix}
\mathbf{x} = \mathbf{y}
$$

をみたす $\mathbf{x} \in \boldsymbol{K}^N$ を求める．

$$
a_{11} \neq 0, \quad a_{22} \neq 0, \quad \cdots, \quad a_{NN} \neq 0
$$

の場合を扱う．まず

$$
x_N = \frac{1}{a_{NN}} y_N \tag{4.7}
$$

でなければならない．また第 $N-1$ 行をみると

$$
a_{N-1,N-1} x_{N-1} + a_{N-1,N} x_N = y_{N-1}
$$

であるから

$$
x_{N-1} = \frac{1}{a_{N-1,N-1}} \left( y_{N-1} - a_{N-1,N} x_N \right)
$$

となっている．これに (4.7) を代入して，

$$
x_{N-1} = \frac{1}{a_{N-1,N-1}} \left( y_{N-1} - \frac{a_{N-1,N}}{a_{NN}} y_N \right)
$$

が得られる．このようにして $x_N, \cdots, x_{j+1}$ が求められるが，さらに第 $j$ 行から

$$
x_j = \frac{1}{a_{jj}} \left( y_j - \sum_{k=j+1}^{N} a_{jk} x_k \right), \quad j = N-1, \cdots, 1
$$

が導かれ，これに $x_N, \cdots, x_{j+1}$ を代入して $x_j$ を得ることができる．このように $x_N$ から初めて $x_{N-1}, x_{N-2}, \cdots$ と順次解を求めていく方法を**後退代入法**という．

ところで，同様の方法で，下三角行列の場合でも，与えられた $\mathbf{y} \in \boldsymbol{K}^N$ に対して

$$\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{N-1,N-1} & 0 \\ a_{N1} & a_{N2} & \cdots & a_{N,N-1} & a_{NN} \end{pmatrix} \mathbf{x} = \mathbf{y}$$

をみたす $\mathbf{x} \in \boldsymbol{K}^N$ が求められる．ただし $a_{11} \neq 0, a_{22} \neq 0, \cdots, a_{NN} \neq 0$ とする．このとき，まず

$$x_1 = \frac{1}{a_{11}} y_1$$

が得られる．一般に第 $j$ 行から

$$x_j = \frac{1}{a_{jj}} \left( y_j - \sum_{k=1}^{j-1} a_{jk} x_k \right), \quad j = 2, \cdots, N$$

を得るが，これを用いて $x_1$ から始めて $x_2, x_3, \cdots$ と順次求めることができる．この方法を**前進代入法**という．

**例題 4.1**　次の連立 1 次方程式を解け．

$$\begin{cases} x_1 + 2x_2 = 1 \\ x_2 + \phantom{2}x_3 = 0 \\ x_1 + \phantom{2}x_3 = 1 \end{cases}$$

**答．**

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{array}\right) \xrightarrow[R_2 \leftrightarrow R_3]{} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[R_2 - R_1]{} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \xrightarrow[R_3 + \frac{1}{2} R_2]{} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & \frac{3}{2} & 0 \end{array}\right).$$

後は後退代入法により $x_3 = 0, x_2 = 0, x_1 = 1 - 2x_2 = 1$． ■

## 4.3　基本行列の積による行列の変形

ガウスの消去法では，次の三つのタイプの変形が方程式系の各式に課せられた．

(a)　ある行をスカラー倍する．

(b)　ある行をスカラー倍して別の行に加える．

(c)　ある行と別の行を入れ替える．

これらの変形を**行に関する基本変形**あるいは**行に関する初等変形**という．行列の基本変形は線形代数ではよく利用される操作である．後で学ぶ行列式の計算，$LU$ 分解，行列標準化などで使われる．本節では基本変形が後述の基本行列と呼ばれる行列との積によって実現できることを示す．これがどのようなことであるかを示すために，まず $(3, N)$ 行列の場合を例にとって述べる．

### 4.3.1　具体例

**基本変形 (a) を実行する行列．**

$(3, N)$ 行列

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix}$$

を変形することを考える．たとえば第 2 行を $\lambda$ 倍するという基本変形は

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \lambda a_{21} & \cdots & \lambda a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix}$$

であるが，この変形自身は行列

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.8}$$

を左からかけることにより実行される．実際

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \lambda a_{21} & \cdots & \lambda a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix}$$

である．したがって連立 1 次方程式

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

の第 2 行を $\lambda$ 倍するという操作は，この方程式の両辺に左側から (4.8) をかけることにより，次のように実行できる．

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

ここで

$$(\text{左辺}) = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \lambda a_{21} & \cdots & \lambda a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

$$(\text{右辺}) = \begin{pmatrix} b_1 \\ \lambda b_2 \\ b_3 \end{pmatrix}.$$

**基本変形 (b) を実行する行列.**

次に

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix}$$

の第 2 行の $\lambda$ 倍を第 3 行に加えるという基本変形を考えてみよう．結果は

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} + \lambda a_{21} & \cdots & a_{3N} + \lambda a_{2N} \end{pmatrix}$$

である．この変形は行列

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix} \tag{4.9}$$

を左からかけることにより，次のように実行できる．

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ \lambda a_{21} + a_{31} & \cdots & \lambda a_{2N} + a_{3N} \end{pmatrix}.$$

したがって方程式

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

の第 2 行の $\lambda$ 倍を第 3 行に加えるという操作は，この方程式の両辺に左側から (4.9) をかけることにより，次のようにできる．

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31}+\lambda a_{21} & \cdots & a_{3N}+\lambda a_{2N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3+\lambda b_2 \end{pmatrix}.$$

**基本変形 (c) を実行する行列．**

第 2 行と第 3 行を入れ替えるという操作は

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{4.10}$$

を左からかけることにより，

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{31} & \cdots & a_{3N} \\ a_{21} & \cdots & a_{2N} \end{pmatrix}$$

と実行できる．したがって方程式

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

の第 2 行と第 3 行を入れ替えるという操作は，両辺に左側から (4.10) をかけることにより，

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{31} & \cdots & a_{3N} \\ a_{21} & \cdots & a_{2N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2N} \\ a_{31} & \cdots & a_{3N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_3 \\ b_2 \end{pmatrix}$$

と実行される.

### 4.3.2 一般の場合

一般の場合，基本変形を実行する行列は次のように定義される.

$E_i^{(N)}(\lambda)$ を，対角成分のうち $(i,i)$ 成分が $\lambda$ で，それ以外の対角成分は 1，かつ対角成分以外の成分はすべて 0 となっている $N$ 次正方行列とする．すなわち

$$E_i^{(N)}(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{4.11}$$

である．ただし，ここで $O$ は表示されたところの成分がすべて 0 であることを表す略記のための記号である．第 4.3.1 節で示した計算と同様にして，容易に次のことがわかる.

> **$E_i^{(N)}(\lambda)$ を $A \in \boldsymbol{K}^{N\times M}$ に左からかけると，$A$ の第 $i$ 行を $\lambda$ 倍したものになる．すなわち $\lambda R_i$ の操作が実行される.**

$i \neq j$ のとき $E_{i,j}^{(N)}(\lambda)$ を $(i,j)$ 成分が $\lambda$，対角成分が 1，それ以外の成分はすべ

て0であるような $N$ 次正方行列とする．たとえば $i<j$ の場合，

$$E_{i,j}^{(N)}(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & \cdots & \lambda & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{4.12}$$

(ただし，$\lambda$ は第 $i$ 行，第 $j$ 列のところに位置している) である．また，$j<i$ の場合は，

$$E_{i,j}^{(N)}(\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & \lambda & \cdots & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{4.13}$$

(ただし，$\lambda$ は第 $i$ 行，第 $j$ 列のところに位置している) である．このとき第4.3.1節での計算と同様にして，次のことが容易に確認できる．

> **$E_{i,j}^{(N)}(\lambda)$ を $A \in \boldsymbol{K}^{N\times M}$ に左からかけると，$A$ の第 $j$ 行の $\lambda$ 倍を第 $i$ 行に加えたものになる．すなわち $R_i + \lambda R_j$ が実行される．**

$i \neq j$ のとき $F_{i,j}^{(N)}$ を，$(i,i)$ 成分と $(j,j)$ 成分が0であり，それ以外の対角成分は1，かつ $(i,j)$ 成分と $(j,i)$ 成分が1，そしてそれ以外の成分が0となっているような $N$ 次正方行列とする．すなわち

$$F_{i,j}^{(N)} = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & O & \\ & & 1 & & & & & & & \\ & & & 0 & & \cdots & & 1 & & \\ & & & & 1 & & & & & \\ & & & \vdots & & \ddots & & \vdots & & \\ & & & & & & 1 & & & \\ & & & 1 & & \cdots & & 0 & & \\ & & & & & & & & 1 & \\ & O & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix} \tag{4.14}$$

($(i,j)$ 成分と $(j,i)$ 成分は 1 である) とする．このとき次のことが容易にわかる．

> **$F_{i,j}^{(N)}$ を$A \in \boldsymbol{K}^{N\times M}$ に左からかけると，$A$ の第 $i$ 行と第 $j$ 行を入れ替えたものになる．すなわち$R_i \leftrightarrow R_j$ の操作が実行される．**

以上のことから，$(N, M)$ 行列の行に関する基本変形が行列 $E_i^{(N)}(\lambda)$, $E_{i,j}^{(N)}(\lambda)$, $F_{i,j}^{(N)}$ を左側からかけることにより実行できることがわかった．まとめると次のようになる．

**定理 4.1**　$N$ 次行列 $E_i^{(N)}(\lambda)$, $E_{i,j}^{(N)}(\lambda)$, $F_{i,j}^{(N)}$ を $(N, M)$ 行列に左側からかけることにより，次のように行列が変形される．

$$\begin{cases} E_i^{(N)}(\lambda) & \cdots \quad \text{第 } i \text{ 行を}\lambda\text{倍する．} & (\lambda R_i) \\ E_{i,j}^{(N)}(\lambda) & \cdots \quad \text{第 } i \text{ 行に第 } j \text{ 行の}\lambda\text{倍を加える．} & (R_i + \lambda R_j) \\ F_{i,j}^{(N)} & \cdots \quad \text{第 } i \text{ 行と第 } j \text{ 行を入れ替える．} & (R_i \leftrightarrow R_j) \end{cases} \tag{4.15}$$

**定義 4.1**　$E_i^{(N)}(\lambda)$, $E_{i,j}^{(N)}(\lambda)$, $F_{i,j}^{(N)}$ を ($N$ **次**) **基本行列**という．$N$ を明記しなくとも混乱を生じない場合は

$$E_i(\lambda) = E_i^{(N)}(\lambda), \quad E_{i,j}(\lambda) = E_{i,j}^{(N)}(\lambda), \quad F_{i,j} = F_{i,j}^{(N)}$$

と略記する．

**問題 4.1**　(4.15) を確認せよ．$(N, 3)$ 行列 $A$ と 3 次基本行列 $E_2^{(3)}(\lambda)$, $E_{2,3}^{(3)}(\lambda)$,

$F_{2,3}^{(3)}$ を $A$ の右側からかけることにより $A$ がどのように変形されるか？　一般に基本行列を右側からかけるとどのような変形が行われるか？

さて，ここで定義した $E_i^{(N)}(\lambda)$, $E_{i,j}^{(N)}(\lambda)$, $F_{i,j}^{(N)}$ はじつは例 2.4 で定義した一般化基本行列の特別な場合に他ならないことが容易に確認できる．

**定理 4.2**　$\mathbf{e}_1, \cdots, \mathbf{e}_N$ を $N$ 次元基本ベクトルとする ((1.1) 参照)．$\mathbf{e}_i\mathbf{e}_j^T$ は $N$ 次正方行列で，$(i,j)$ 成分が 1 で，それ以外の成分はすべて 0 である．さらに

$$\begin{aligned} E_i^{(N)}(\lambda) &= I_N + (\lambda - 1)\,\mathbf{e}_i\mathbf{e}_i^T, \\ E_{i,j}^{(N)}(\lambda) &= I_N + \lambda\,\mathbf{e}_i\mathbf{e}_j^T, \quad i \neq j, \\ F_{i,j}^{(N)} &= I_N - (\mathbf{e}_i - \mathbf{e}_j)\,(\mathbf{e}_i - \mathbf{e}_j)^T, \quad i \neq j \end{aligned}$$

となっている．

ここでは具体的な $E_2^{(3)}(\lambda)$, $E_{3,2}^{(3)}(\lambda)$, $F_{2,3}^{(3)}$ に対して定理の主張を確認する．一般の場合も同様にして証明できる．

$$\mathbf{e}_2\mathbf{e}_2^T = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

であるから

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_2^{(3)}(\lambda)$$

である．また

$$\mathbf{e}_3\mathbf{e}_2^T = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

より

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \lambda & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda & 1 \end{pmatrix} = E_{3,2}^{(3)}(\lambda)$$

となっている．

$$
\begin{aligned}
&I_3 - (\mathbf{e}_2 - \mathbf{e}_3)(\mathbf{e}_2 - \mathbf{e}_3)^T \\
&= I_3 - \mathbf{e}_2\mathbf{e}_2^T - \mathbf{e}_3\mathbf{e}_3^T + \mathbf{e}_2\mathbf{e}_3^T + \mathbf{e}_3\mathbf{e}_2^T \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = F_{2,3}^{(3)}
\end{aligned}
$$

が成り立つ．

**定理 4.3**　$\lambda \neq 0$ とする．このとき $N$ 次基本行列 $E_i(\lambda)$(定義 4.1 参照) は正則であり，

$$
E_i(\lambda)^{-1} = E_i\left(\frac{1}{\lambda}\right).
$$

また $i \neq j$ の場合，$\lambda \in \boldsymbol{K}$ に対して $E_{i,j}(\lambda)$ は正則，$F_{i,j}$ は正則であり，

$$
\begin{aligned}
E_{i,j}(\lambda)^{-1} &= E_{i,j}(-\lambda) \\
F_{i,j}^{-1} &= F_{i,j} = F_{j,i}
\end{aligned}
$$

である．

**証明.** $E_i(\lambda)$ は第 $i$ 行を $\lambda$ 倍し，$E_i(\lambda^{-1})$ は第 $i$ 行を $\dfrac{1}{\lambda}$ 倍するので，

$$
E_i(\lambda)\,E_i\left(\lambda^{-1}\right) = I_N, \quad E_i\left(\lambda^{-1}\right)E_i(\lambda) = I_N
$$

である．$E_{i,j}(\lambda)$ は第 $i$ 行に第 $j$ 行の $\lambda$ 倍を加え，$E_{i,j}(-\lambda)$ は第 $i$ 行から第 $j$ 行の $\lambda$ 倍を引くので，$E_{i,j}(\lambda)$ の逆行列は $E_{i,j}(-\lambda)$ である．$F_{i,j}$ は第 $i$ 行と第 $j$ 行を入れ替える操作を施すものであるから，$F_{i,j} = F_{j,i}$ かつ $F_{i,j}F_{i,j} = I_N$ である．

■

**問題 4.2**　定理 4.3 を定理 4.2 および例 2.4 で述べた結果を用いて証明せよ．

## 4.4　逆行列の計算――掃き出し法

ガウスの消去法を利用すれば，$N$ 次正則行列 $A = (a_{ij})$ の逆行列を計算することができる．本節ではその方法を述べる．

もし $A$ に対してガウスの消去法の操作が途中で停止することなく最後まで実行できたならば，ある有限個の基本行列 $E_1, \cdots, E_k$ を左側からかけて

$$E_k \cdots E_1 A = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2N}^{(2)} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & a_{NN}^{(N)} \end{pmatrix},$$

$$\text{ただし} \quad a_{11}^{(1)} \cdots a_{NN}^{(N)} \neq 0$$

のように上三角行列に変形できる (じつは $A$ が正則行列ならば常にこのような変形ができることがわかる．証明は後述の問題 5.4 参照)．このとき

$$(\text{第 } j \text{ 行}) - \frac{a_{j\,N}^{(j)}}{a_{NN}^{(N)}} \times (\text{第 } N \text{ 行}), \quad j = N-1, \cdots, 1$$

なる基本変形により

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,N-1}^{(1)} & 0 \\ 0 & a_{22}^{(2)} & \cdots & a_{2,N-1}^{(2)} & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & a_{N-1,N-1}^{(N-1)} & 0 \\ 0 & 0 & \cdots & 0 & a_{NN}^{(N)} \end{pmatrix}$$

とできる．引き続き

$$(\text{第 } j \text{ 行}) - \frac{a_{j\,N-1}^{(j)}}{a_{N-1,N-1}^{(N-1)}} \times (\text{第 } N-1 \text{ 行}), \quad j = N-2, \cdots, 1$$

なる基本変形をしていくと

$$\begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1,N-2}^{(1)} & 0 & 0 \\ & \ddots & \vdots & \vdots & \\ & & a_{N-2,N-2}^{(N-2)} & 0 & \vdots \\ & O & & a_{N-1,N-1}^{(N-1)} & 0 \\ & & & & a_{NN}^{(N)} \end{pmatrix}$$

となり，さらにこの操作を続けていけば

$$\begin{pmatrix} a_{11}^{(1)} & & O \\ & \ddots & \\ O & & a_{NN}^{(N)} \end{pmatrix}$$

という対角行列が得られる．そして各第 $j$ 行を $\dfrac{1}{a_{jj}^{(j)}}$ 倍すると (これも基本変形)

$$\begin{pmatrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{pmatrix}$$

が得られる．

以上の基本変形を基本行列で表したものをたとえば $E_{k+1}, \cdots, E_n$ とおくと

$$E_n \cdots E_1 A = I_N \tag{4.16}$$

が得られる．したがって $B = E_n \cdots E_1$ とおけば $BA = I_N$ となっている．一方

$$AB = I_N$$

でもあることが次のようにして証明できる．定理 4.3 より，$E_j^{-1}$ が存在するので，(4.16) の両辺に左側から $E_1^{-1} \cdots E_n^{-1}$ をかければ

$$A = E_1^{-1} \cdots E_n^{-1} E_n \cdots E_1 A = E_1^{-1} \cdots E_n^{-1} I_N = E_1^{-1} \cdots E_n^{-1}.$$

したがって

$$AB = AE_n \cdots E_1 = E_1^{-1} \cdots E_n^{-1} E_n \cdots E_1 = I_N$$

となっている．以上より $A^{-1} = B$ であることがわかる．

ここで

$$A^{-1} = B = E_n \cdots E_1 I_N$$

が成り立つが，このことは $A$ の逆行列 $A^{-1}$ は $A$ を $I_N$ に変形するための基本変形と同じ操作を $I_N$ に施すことによって得られることを示している．

具体例で検証してみよう．

**例 4.3**

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 4 & -3 \end{pmatrix}$$

とする．$A$ を単位行列に基本変形していくのだが，すでに見たように逆行列は同じ基本変形を $I_3$ に対して施すと得られるので，$A$ と同時に $A$ の変形と同様の変形を $I_3$ に対しても行う．まず $A$ の第 1 行の 2 倍を第 2 行，第 3 行から引くと

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 4 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix}$$

である．これと同じ操作を $I_3$ に施すと

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

これを次のように略記する．

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 2 & 4 & -3 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[R_2-2R_1,\ R_3-2R_1]{} \left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -2 & 1 & 0 \\ 0 & 2 & -1 & -2 & 0 & 1 \end{array}\right).$$

さらに次のように変形する．

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -2 & 1 & 0 \\ 0 & 2 & -1 & -2 & 0 & 1 \end{array}\right) \xrightarrow[R_3+2R_2]{} \left(\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -6 & 2 & 1 \end{array}\right).$$

これで $A$ は上三角行列に変形された．ここからさらに (第 2 行)$-$(第 3 行) , (第 1 行)$+$(第 3 行) をすると

$$\xrightarrow[R_2-R_3,\ R_1+R_3]{} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & -5 & 2 & 1 \\ 0 & -1 & 0 & 4 & -1 & -1 \\ 0 & 0 & 1 & -6 & 2 & 1 \end{array}\right)$$

である．次に (第 1 行) $+$ (第 2 行) として

$$\xrightarrow[R_1+R_2]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 4 & -1 & -1 \\ 0 & 0 & 1 & -6 & 2 & 1 \end{array}\right)$$

となる．最後に第 2 行に $-1$ をかければ,

$$\xrightarrow[-R_2]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -4 & 1 & 1 \\ 0 & 0 & 1 & -6 & 2 & 1 \end{array}\right)$$

となり，逆行列

$$A^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 1 & 1 \\ -6 & 2 & 1 \end{pmatrix}$$

が得られる．

**例題 4.2** $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ の逆行列を求めよ．

**答．** 第 1 行と第 2 行を入れ替えることから始める．

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_1 \leftrightarrow R_2]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[R_1 - R_3]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

ゆえに求める逆行列は $\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ である．■

**問題 4.3** 次の行列の逆行列を求めよ．

(1) $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (2) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

**問題 4.4** $\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & -1 & 0 & 1 \\ -1 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ が $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$ の逆行列であることを直接計算して確かめよ.

**問題 4.5** 　次の行列の逆行列を求めよ.

(1) $\begin{pmatrix} 1 & 3 \\ 5 & 2 \end{pmatrix}$ 　　(2) $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 1 & 1 \end{pmatrix}$

# 第5章　行列式

$N$ 次正方行列の正則性を判定し，さらに逆行列を記述できる量に行列式というものがある．この章では $N$ 次正方行列に対する行列式を定義し，それに関する基本的な性質を述べる．まず 2 次正方行列の場合に行列式の定義をする．そのために第 2.4 節の例 2.3 で述べた 2 次正方行列の逆行列に関する結果を改めて述べておく．

**定理 5.1**　$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \boldsymbol{K}^{2\times 2}$ とする．$A$ が正則であるための必要十分条件は

$$ad - bc \neq 0$$

となることである．$ad - bc \neq 0$ のとき $A$ の逆行列は

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

である．

ここで $A$ が正則であるための条件として記された量 $ad - bc$ に着目して，次のように $A$ の行列式を定義する．

**定義 5.1**　行列 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \boldsymbol{K}^{2\times 2}$ に対して，

$$|A| = ad - bc$$

を $A$ の**行列式**という．

この定義を用いれば，$A$ が正則であるための必要十分条件は $|A| \neq 0$ をみたすことと表すことができ，また $A$ の逆行列の記述も行列式を用いて可能になる．

本章の目的はこの行列式の概念を一般の $N$ 次正方行列の場合に一般化することである．その一般化した行列式を用いることにより，$N$ 次正方行列が正則であること，そして正則な行列の逆行列を求めることができる．

**ポイント5**　$N$ 次正方行列の行列式の定義 (第5.3節参照) は少しわかりにくいかもしれないが，具体的な計算練習などを行って慣れておくとよいだろう．なお，次章で学ぶ余因子展開を用いる行列式の計算方法も便利である．

## 5.1　置換

一般の $N$ 次正方行列の行列式の定義に必要なので，本節と次節では置換および置換の符号の定義をする．

集合 $\{1,2,\cdots,N\}$ から集合 $\{1,2,\cdots,N\}$ への全単射 $\sigma$ を $\{1,2,\cdots,N\}$ の $N$ **次置換**あるいは単に**置換**という．定義から置換 $\sigma$ は次の性質をみたす．

$$i \neq j \quad \text{ならば} \quad \sigma(i) \neq \sigma(j), \tag{5.1}$$

$$\{\sigma(1),\cdots,\sigma(N)\} = \{1,\cdots,N\}. \tag{5.2}$$

これより $\sigma(1),\cdots,\sigma(N)$ は $1,\cdots,N$ を並べ替えたものにほかならないことがわかる．

$\{1,\cdots,N\}$ の置換全体のなす集合を $S_n$ と表す．$1,\cdots,N$ の並べ替え方は $N!$ 通りあるので，$\{1,2,\cdots,N\}$ の置換は全部で $N!$ 個ある．

いま $1,2,\cdots,N$ を並べ替えた数 $i_1,\cdots,i_N$ を考える．$N$ 次置換 $\sigma$ が $\sigma(1) = i_1$, $\sigma(2) = i_2$, $\cdots$, $\sigma(N) = i_N$ をみたすとき，$\sigma$ を

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix}$$

と表す．たとえば $\{1,2,3\}$ の置換 $\sigma$ が $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$ をみたすときは

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

である．

さて，$\{1,2,3\}$ の置換は $3! = 6$ 個あるが，それらはこの記法を用いれば

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\quad \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\quad \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\quad \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\quad \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

と表すことができる．なお

$$\sigma = \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

は $\sigma$ が $1 \to 3,\ 2 \to 2,\ 3 \to 1$ なる変換をしていることを意味しているので，置換の上側に並んでいる数字 $1, 2, 3$ の順序を変えても，それに対応した数字を下側に並べれば，それも同じ置換を表すものとする．たとえば

$$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix} = \begin{pmatrix}1&3&2\\3&1&2\end{pmatrix} = \begin{pmatrix}2&1&3\\2&3&1\end{pmatrix}$$

のように表してもよい．ただし置換の上側の数字と下側の数字の対応関係は崩してはいけない．

特に置換 $\begin{pmatrix}1&2&\cdots&N\\1&2&\cdots&N\end{pmatrix}$ を**恒等置換**といい，$1_N$ と表す．

二つの置換 $\sigma, \tau$ の合成 $\tau \circ \sigma$ を $\tau\sigma$ と表し，これを置換 $\sigma, \tau$ の**積**という．たとえば

$$\sigma = \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\quad \tau = \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

のとき $\tau\sigma$ は次のような写像になる．

$$\begin{aligned}1 &\stackrel{\sigma}{\longmapsto} 2 \stackrel{\tau}{\longmapsto} 2\\ 2 &\stackrel{\sigma}{\longmapsto} 3 \stackrel{\tau}{\longmapsto} 1\\ 3 &\stackrel{\sigma}{\longmapsto} 1 \stackrel{\tau}{\longmapsto} 3\end{aligned}$$

ここで $i \stackrel{\sigma}{\longmapsto} j$ は $\sigma(i) = j$ を意味し，$i \stackrel{\tau}{\longmapsto} j$ は $\tau(i) = j$ を意味する．

$$\tau\sigma = \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$$

である．

**例 5.1**　$\sigma = \begin{pmatrix}1&2&\cdots&N\\i_1&i_2&\cdots&i_N\end{pmatrix}, \tau = \begin{pmatrix}i_1&i_2&\cdots&i_N\\j_1&j_2&\cdots&j_N\end{pmatrix}$ に対して

$$
\begin{aligned}
\tau\sigma &= \begin{pmatrix} i_1 & i_2 & \cdots & i_N \\ j_1 & j_2 & \cdots & j_N \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & \cdots & N \\ j_1 & j_2 & \cdots & j_N \end{pmatrix}
\end{aligned}
$$

である．

**例 5.2**　$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ とする．

$\tau = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$ と表せるので，

$$
\tau\sigma = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}
$$

である．また $\sigma = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}$ であるから

$$
\sigma\tau = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}
$$

である．したがって

$$
\sigma\tau \neq \tau\sigma.
$$

**命題 5.2**　$\sigma = \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix}$ に対して，

$\sigma^{-1} = \begin{pmatrix} i_1 & i_2 & \cdots & i_N \\ 1 & 2 & \cdots & N \end{pmatrix}$ とおく．このとき，

$$
\sigma^{-1}\sigma = \sigma\sigma^{-1} = 1_N.
$$

**証明.**

$$
\begin{aligned}
\sigma^{-1}\sigma &= \begin{pmatrix} i_1 & i_2 & \cdots & i_N \\ 1 & 2 & \cdots & N \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & \cdots & N \\ 1 & 2 & \cdots & N \end{pmatrix} = 1_N.
\end{aligned}
$$

$$\sigma\sigma^{-1} = \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix} \begin{pmatrix} i_1 & i_2 & \cdots & i_N \\ 1 & 2 & \cdots & N \end{pmatrix}$$
$$= \begin{pmatrix} i_1 & i_2 & \cdots & i_N \\ i_1 & i_2 & \cdots & i_N \end{pmatrix} = 1_N.$$

■

**定義 5.2**　$\sigma^{-1}$ を $\sigma$ の**逆置換**という．

ところで，置換 $\sigma$ が $\sigma(j) = j$ となる $j$ をもつとき，すなわち順序を動かさない部分をもつとき，この部分を省略して次のように表すことがある．

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

特に二つの文字 $i, j$ を $j, i$ と入れ替え，他の文字は動かさないとき，これを単に $(i, j)$ と表し，**互換**という．すなわち

$$\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & N \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & N \end{pmatrix} = (i, j)$$

である．たとえば

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2, 3), \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = (1, 3)$$

と表す．

置換が与えられたとき，次のことが成り立つ．

**定理 5.3**　任意の置換は互換の積で表せる．ただし，恒等置換は 0 個の互換の積，互換は 1 個の互換の積と考える．

**証明.** $\sigma$ を $n$ 文字の置換とする．$n$ に関する帰納法で示す．

$n = 1, 2$ の場合は明らかである．$n - 1 \geq 2$ で成立するとして $n$ の場合を示す．

$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ とする．

$i_n = n$ の場合は，$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 \\ i_1 & i_2 & \cdots & i_{n-1} \end{pmatrix}$ であるから，帰納法の仮定より $\sigma$ は互換の積で表せる．

$i_n \neq n$ の場合は，$\tau = (i_n, n)$ とすると，

$$\tau\sigma = (i_n, n)\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ j_1 & j_2 & \cdots & j_{n-1} & n \end{pmatrix}$$

であるから，$\tau\sigma$ は $n-1$ 文字の置換とみなすことができる．したがって帰納法の仮定から $\tau\sigma$ はある互換 $\tau_1, \cdots, \tau_k$ の積として

$$\tau\sigma = \tau_1 \cdots \tau_k$$

と表せる．$\tau\tau = 1_n$, $1_n\sigma = \sigma$ であるから (問題 5.2 参照)

$$\sigma = \tau\tau\sigma = \tau\tau_1 \cdots \tau_k$$

を得る．■

一般に，置換の互換の積への分解は一意的ではない．たとえば

$$(4,3)(2,3)(1,3) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} = (1,2)(3,2)(4,2).$$

**問題 5.1**　$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ とする．このとき

$$\tau\sigma, \quad \sigma\tau, \quad \sigma\tau\sigma^{-1}, \quad \tau\sigma\tau^{-1}$$

を求めよ．

**問題 5.2**　$N$ 次置換が互換 $\sigma = (i, j)$ の場合，$\sigma\sigma = 1_N$ であることを示せ．

## 5.2　置換の符号と偶置換，奇置換

置換 $\sigma$ に対して，その符号と呼ばれる量 $\mathrm{sgn}\,(\sigma)$ を定義する．そのため，$x_1, \cdots, x_n$ を変数とする式

$$\Delta(x_1, \cdots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i)$$

を考える．ただし，ここで右辺は $1 \le i < j \le n$ をみたすすべての $i, j$ にわたって $x_j - x_i$ の積をとることを表している．すなわち

$$\Delta(x_1,\cdots,x_n)=\left\{\begin{array}{cccc}(x_2-x_1) & (x_3-x_1) & \cdots & (x_n-x_1)\\ & \times(x_3-x_2) & \cdots & (x_n-x_2)\\ & & \ddots & \vdots\\ & & & \times(x_n-x_{n-1})\end{array}\right.$$

である．これを**差積**という．混乱のないときは $\Delta(x_1,\cdots,x_n)$ を単に $\Delta$ と表す．$n$ 次置換 $\sigma$ に対して

$$\sigma\Delta(x_1,\cdots,x_n)=\prod_{i<j}(x_{\sigma(j)}-x_{\sigma(i)})$$

$$=\left\{\begin{array}{cccc}(x_{\sigma(2)}-x_{\sigma(1)}) & (x_{\sigma(3)}-x_{\sigma(1)}) & \cdots & (x_{\sigma(n)}-x_{\sigma(1)})\\ & \times(x_{\sigma(3)}-x_{\sigma(2)}) & \cdots & (x_{\sigma(n)}-x_{\sigma(2)})\\ & & \ddots & \vdots\\ & & & \times(x_{\sigma(n)}-x_{\sigma(n-1)})\end{array}\right.$$

と定義する．定義より

$$\sigma\Delta(x_1,\cdots,x_n)=\Delta(x_1,\cdots,x_n)$$

または

$$\sigma\Delta(x_1,\cdots,x_n)=-\Delta(x_1,\cdots,x_n)$$

のどちらかが成り立つことがわかる．

**例 5.3** $\sigma=\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 1 & 3\end{pmatrix}$ に対して，

$$\sigma\Delta(x_1,\cdots,x_4)=\left\{\begin{array}{ccc}(x_4-x_2) & (x_1-x_2) & (x_3-x_2)\\ & \times(x_1-x_4) & (x_3-x_4)\\ & & \times(x_3-x_1)\end{array}\right.$$

$$=\left\{\begin{array}{ccc}(x_4-x_2) & (-1)(x_2-x_1) & (x_3-x_2)\\ & \times(-1)(x_4-x_1) & (-1)(x_4-x_3)\\ & & \times(x_3-x_1)\end{array}\right.$$

$$=-\Delta(x_1,\cdots,x_4).$$

また $\sigma=\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 1 & 3 & 2\end{pmatrix}$ に対して

$$
\begin{aligned}
\sigma\Delta(x_1,\cdots,x_4) &= \left\{\begin{array}{lll}
(x_1-x_4) & (x_3-x_4) & (x_2-x_4) \\
 & \times(x_3-x_1) & (x_2-x_1) \\
 & & \times(x_2-x_3)
\end{array}\right. \\
&= \left\{\begin{array}{lll}
(-1)(x_4-x_1) & (-1)(x_4-x_3) & (-1)(x_4-x_2) \\
 & \times(x_3-x_1) & (x_2-x_1) \\
 & & \times(-1)(x_3-x_2)
\end{array}\right. \\
&= \Delta(x_1,\cdots,x_4).
\end{aligned}
$$

置換に対してその符号という量が次のように定義される．

**定義 5.3**　$\sigma$ を $N$ 次置換とする．$x_1,\cdots,x_n$ の差積 $\Delta$ に対して，

$$
\left\{\begin{array}{ll}
\sigma\Delta=\Delta \text{ をみたすとき} & \operatorname{sgn}(\sigma)=1 \\
\sigma\Delta=-\Delta \text{ をみたすとき} & \operatorname{sgn}(\sigma)=-1
\end{array}\right.
$$

と定める．$\operatorname{sgn}(\sigma)$ を $\sigma$ の**符号**(sign) という．特に $\operatorname{sgn}(\sigma)=1$ なるとき，$\sigma$ を**偶置換**という．$\operatorname{sgn}(\sigma)=-1$ となるとき，$\sigma$ を**奇置換**という．

たとえば例 5.3 で挙げた $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$ は奇置換，$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ は偶置換である．

符号の定義から容易に次のことがわかる．$\sigma$ が偶置換か奇置換であるかは，

$$
\begin{array}{llll}
(\sigma(2),\sigma(1)), & (\sigma(3),\sigma(1)), & \cdots, & (\sigma(n),\sigma(1)) \\
 & (\sigma(3),\sigma(2)), & \cdots, & (\sigma(n),\sigma(2)) \\
 & & \ddots & \vdots \\
 & & & (\sigma(n),\sigma(n-1))
\end{array}
$$

という数の組の中で，$(\sigma(p),\sigma(q))$ のうち $\sigma(p)<\sigma(q)$ であるようなものが偶数個あれば偶置換，奇数個あれば奇置換である (たとえば例 5.3 でこの主張が正しいことを確認せよ)．符号が次のようにも表せることも容易にわかる．

**定理 5.4**　$\sigma$ を $\{1,\cdots,N\}$ の置換とする．このとき，

$$
\operatorname{sgn}(\sigma)=\prod_{1\le i<j\le N}\frac{\sigma(j)-\sigma(i)}{j-i}
$$

である．

**定理 5.5**　$\sigma$ を $\{1,\cdots,N\}$ の置換とし，$\tau$ を互換とする．このとき，

$$(\tau\sigma)\,\Delta = -\sigma\Delta$$

である．すなわち，$\sigma$ が偶置換なら $\tau\sigma$ は奇置換であり，$\sigma$ が奇置換ならば $\tau\sigma$ は偶置換である．

**証明.** $\sigma = \begin{pmatrix} 1 & \cdots & N \\ p_1 & \cdots & p_N \end{pmatrix}$ とする．$\sigma\Delta = \prod_{i<j}(x_{p_j} - x_{p_i})$ である．互換を $\tau = (i,k)\,(i<k)$ とすると，$\tau$ により，次の置き換えがなされる．

$$\left\{\ x_{p_k} - x_{p_i} \quad \longrightarrow \quad x_{p_i} - x_{p_k}. \right. \tag{5.3}$$

このほか $h \neq i,\, h \neq k$ なる $h$ に関して次の置き換えがなされる．

$$\begin{cases} x_{p_i} - x_{p_h} & \longrightarrow \quad x_{p_k} - x_{p_h} \\ x_{p_k} - x_{p_h} & \longrightarrow \quad x_{p_i} - x_{p_h} \end{cases} \qquad (h<i<k) \tag{5.4}$$

$$\begin{cases} x_{p_h} - x_{p_i} & \longrightarrow \quad x_{p_h} - x_{p_k} \\ x_{p_k} - x_{p_h} & \longrightarrow \quad x_{p_i} - x_{p_h} \end{cases} \qquad (i<h<k) \tag{5.5}$$

$$\begin{cases} x_{p_h} - x_{p_i} & \longrightarrow \quad x_{p_h} - x_{p_k} \\ x_{p_h} - x_{p_k} & \longrightarrow \quad x_{p_h} - x_{p_i} \end{cases} \qquad (i<k<h) \tag{5.6}$$

ここで (5.3) の入れ替えで符号が一回変わる．また (5.5) で二回符号が変わり，また (5.4), (5.6) では符号の変化はない．結局符号は一回変わるだけである．■

前節で置換は互換の積に分解できることを証明した (定理 5.3 参照)．その分解の仕方は一意的ではないが，次のことがわかる．

**系 5.6**　置換を互換の積に分解したとき，偶置換ならば互換の数は偶数個，奇置換ならば互換の数は奇数個でなければならない．

**証明.** $\sigma$ を置換とし，$\sigma = \tau_1 \cdots \tau_k$ とすると，

$$\begin{aligned} \sigma\Delta &= \tau_1 \cdots \tau_k \Delta = -\tau_2 \cdots \tau_k \Delta = \tau_3 \cdots \tau_k \Delta \\ &= \cdots = (-1)^k \Delta \end{aligned}$$

である．したがって，$\sigma$ が遇置換ならば $k$ は偶数，$\sigma$ が奇置換ならば $k$ は奇数でなければならない．■

**系 5.7**　$\sigma, \tau \in S_N$ に対して

$$\mathrm{sgn}\,(\sigma\tau) = \mathrm{sgn}\,(\sigma)\;\mathrm{sgn}\,(\tau)\,.$$

**証明.** $\sigma, \tau$ をそれぞれ互換の積に分解すると，$\sigma\tau$ はそれらの互換の積で表すことができる．したがって系が示される．■

## 5.3　行列式

準備が終わったので，一般の $N$ 次正方行列に対して行列式を定義する．$N$ 次置換は全部で $N!$ 個あるので，それを $\sigma_1, \cdots, \sigma_{N!}$ とおく．

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$$

に対して，

$$|A| = \sum_{j=1}^{N!} \mathrm{sgn}(\sigma_j) a_{1\sigma_j(1)} a_{2\sigma_j(2)} \cdots a_{N\sigma_j(N)} \tag{5.7}$$

と定義し，これを $A$ の**行列式**という．$|A|$ を $\det A$ あるいは $\det(A)$ とも表す．また

$$\sum_{j=1}^{N!} \mathrm{sgn}(\sigma_j)\, a_{1\sigma_j(1)} \cdots a_{N\sigma_j(N)}$$

を

$$\sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{N\sigma(N)}$$

と表す．これは $N$ 次置換全体のなす集合 $S_N$ の要素全部にわたって和をとることを意味した記法である．

**注意 5.1**　行列式を得るための操作は次のように言うこともできる．各行から成分を一つずつ，しかも同じ列からは二つ以上とらないように選び，それを $a_{1\,i_1}, \cdots, a_{N\,i_N}$ とする．そして $\begin{pmatrix} 1 & \cdots & N \\ i_1 & \cdots & i_N \end{pmatrix}$ が偶置換ならば $a_{1\,i_1} \cdots a_{N\,i_N}$，また奇置換ならば $-a_{1\,i_1} \cdots a_{N\,i_N}$ とする．$i_1, \cdots, i_N$ の選び方は全部で $N!$ 個あるが，それらをすべて足し合わせたものが $|A|$ である．

行列式の定義を 2 次正方行列

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

に当てはめると，

$$\begin{aligned}|A| &= \sum_{\sigma \in S_2} \operatorname{sgn}(\sigma)\ a_{1\sigma(1)} a_{2\sigma(2)} \\ &= \operatorname{sgn}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} a_{11}a_{22} + \operatorname{sgn}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} a_{12}a_{21} \\ &= a_{11}a_{22} - a_{12}a_{21}\end{aligned}$$

となり，すでに述べた行列式の定義と一致する．3 次正方行列

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

の場合は

$$\begin{aligned}|A| &= \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \\ &= \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} a_{11}a_{22}a_{33} + \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} a_{11}a_{23}a_{32} \\ &\quad + \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} a_{12}a_{21}a_{33} + \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} a_{12}a_{23}a_{31} \\ &\quad + \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} a_{13}a_{21}a_{32} + \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} a_{13}a_{22}a_{31} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} \\ &\quad + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \qquad (5.8)\end{aligned}$$

となる．

行列式は次のように表すこともできる．

**定理 5.8**　$A = (a_{jk})_{\substack{j=1,\cdots,N \\ k=1,\cdots,N}}$ に対して，

$$|A| = \sum_{\tau \in S_N} \operatorname{sgn}(\tau)\, a_{\tau(1)1} a_{\tau(2)2} \cdots a_{\tau(N)N}.$$

**証明.** $\sigma$ の逆置換 $\sigma^{-1} = \begin{pmatrix} \sigma(1) & \cdots & \sigma(N) \\ 1 & \cdots & N \end{pmatrix}$ を並べ替えたものを

$$\begin{pmatrix} 1 & \cdots & N \\ \sigma^{-1}(1) & \cdots & \sigma^{-1}(N) \end{pmatrix}$$

とおく．このとき

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{N\sigma(N)} = a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(N)N}$$

である．$\sigma^{-1}\sigma = 1_N$ であり，$1_N$ は偶置換であるから，$\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$ である．ゆえに

$$\mathrm{sgn}(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{N\sigma(N)} = \mathrm{sgn}(\sigma^{-1})\, a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(N)N}$$

となっている．したがって

$$\begin{aligned} |A| &= \sum_{\sigma\in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{N\sigma(N)} \\ &= \sum_{\sigma\in S_N} \mathrm{sgn}(\sigma^{-1})\, a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(N)N} \end{aligned}$$

が成り立つ．

ここで $S_N = \{\sigma^{-1} : \sigma \in S_N\}$ であることを示しておく．$S'_N = \{\sigma^{-1} : \sigma \in S_N\}$ とおく．$\sigma \in S_N$ ならば $\sigma^{-1}$ も $N$ 次置換であるので，$\sigma^{-1} \in S_N$ である．ゆえに $S'_N \subset S_N$ がわかる．また $\sigma \in S_N$ ならば $\sigma = (\sigma^{-1})^{-1}$ であり，$(\sigma^{-1})^{-1} \in S'_N$ であるから $\sigma \in S'_N$ が得られる．ゆえに $S_N \subset S'_N$ も成り立つ．したがって $S_N = S'_N$ である．

よって

$$\begin{aligned} &\sum_{\sigma\in S_N} \mathrm{sgn}(\sigma^{-1})\, a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(N)N} \\ &\quad = \sum_{\tau\in S'_N} \mathrm{sgn}(\tau)\, a_{\tau(1)1}a_{\tau(2)2}\cdots a_{\tau(N)N} \\ &\quad = \sum_{\tau\in S_N} \mathrm{sgn}(\tau)\, a_{\tau(1)1}a_{\tau(2)2}\cdots a_{\tau(N)N} \end{aligned}$$

であることがわかり，定理が証明された．■

2 次正方行列に対する行列式の定義の $N$ 次正方行列の場合への一般化を (5.7) のようにしておくと，2 次の場合と同じく

$$N \text{ 次正方行列 } A \text{ が正則} \Longleftrightarrow |A| \neq 0$$

が成り立つ．ただし，このことの証明は，もう少し行列式の基本的な性質を学んだ後で，$A^{-1}$ を求める公式とともに証明する (定理 6.3 参照).

**例題 5.1**　次の行列式を求めよ．

(1) $\begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 3 \\ 3 & 2 & 4 \end{pmatrix}$　　(2) $\begin{pmatrix} y & z & x \\ x & y & z \\ z & x & y \end{pmatrix}$

**答．** (5.8) より，(1) $-1$,　(2) $x^3 + y^3 + z^3 - 3xyz$. ■

**例題 5.2**　次の等式を証明せよ．

(1) $$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{NN} \end{vmatrix} = a_{11}a_{22}\cdots a_{NN}.$$

(2) $$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & a_{22} & \cdots & a_{2N} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{NN} \end{vmatrix} = a_{11}a_{22}\cdots a_{NN}.$$

(3) $$\begin{vmatrix} a_{11} & \cdots & a_{1,N-1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{N-1,1} & \cdots & a_{N-1,N-1} & 0 \\ a_{N1} & \cdots & a_{N,N-1} & a_{NN} \end{vmatrix} = a_{NN} \begin{vmatrix} a_{11} & \cdots & a_{1,N-1} \\ \vdots & \ddots & \vdots \\ a_{N-1,1} & \cdots & a_{N-1,N-1} \end{vmatrix}.$$

**答．** (1)　$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{N\sigma(N)}$ は $\sigma(j) \neq j$ のときは 0 であるから,

$$\sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{N\sigma(N)} = a_{11}a_{22}\cdots a_{NN}$$

である．

(2)　$N$ 次置換 $\sigma$ が $\sigma(j) \neq j$ となる $j$ をもつとする．$\sigma(j) < j$ の場合 $a_{j\sigma(j)} = 0$ であるから，$a_{1\sigma(1)} \cdots a_{N\sigma(N)} = 0$ である．$\sigma(j) > j$ の場合は，$\sigma(k_0) < k_0$ をみたす $k_0$ が少なくとも一つ存在する．なぜならば，もしもすべての $k$ に対して，$\sigma(k) \geq k$ であれば，

$$\begin{aligned}\sum_{k=1}^{N} \sigma(k) &= \sigma(j) + \sum_{\substack{k=1\\k\neq j}}^{N} \sigma(k) \geq \sigma(j) + \sum_{\substack{k=1\\k\neq j}}^{N} k \\ &> j + \sum_{\substack{k=1\\k\neq j}}^{N} k = \sum_{n=1}^{N} n\end{aligned}$$

であるが，$\sigma$ は $\{1, \cdots, N\}$ の並べ替えであるから

$$\sum_{n=1}^{N} n = \sum_{k=1}^{N} \sigma(k)$$

が成り立っていなければならず，矛盾が導かれる．$\sigma(k_0) < k_0$ なる $k_0$ に対して $a_{k_0\sigma(k_0)} = 0$ であるから，$\sigma(j) > j$ の場合も $a_{1\sigma(1)} \cdots a_{N\sigma(N)} = 0$ となっている．ゆえに

$$\sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{N\sigma(N)} = a_{11}a_{22} \cdots a_{NN}$$

である．

(3)　$\sigma(N) \neq N$ のとき $a_{\sigma(N)N} = 0$ であるから，

$$\sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(N)N} = a_{NN} \sum_{\substack{\sigma \in S_N\\ \sigma(N)=N}} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(N-1),N-1}.$$

さらに

$$\begin{aligned}\sum_{\substack{\sigma \in S_N\\ \sigma(N)=N}} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(N-1),N-1} &= \sum_{\tau \in S_{N-1}} \mathrm{sgn}(\tau)\, a_{\tau(1)1} \cdots a_{\tau(N-1),N-1} \\ &= \begin{vmatrix} a_{11} & \cdots & a_{1,N-1} \\ \vdots & \ddots & \vdots \\ a_{N-1,1} & \cdots & a_{N-1,N-1} \end{vmatrix}\end{aligned}$$

となることがわかる．■

## 5.4 行列式の基本的な性質

行列式を計算する際に，定義に基づいて直接計算するのはたいへんなことが多い．本節では行列式の計算に便利ないくつかの公式を述べる．

### 5.4.1 行に関する変形

**(I)**

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ ta_{k1} & ta_{k2} & \cdots & ta_{kN} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix} = t \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kN} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{vmatrix}.$$

**証明.**

$$\begin{aligned}(\text{左辺}) &= \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots (ta_{k\sigma(k)}) \cdots a_{N\sigma(N)} \\ &= t \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{N\sigma(N)} = (\text{右辺}).\end{aligned}$$

■

**(II)**

$$\begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \cdots & \vdots \\ a_{k1} + a'_{k1} & \cdots & a_{kN} + a'_{kN} \\ \vdots & \cdots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a'_{k1} & \cdots & a'_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}.$$

**証明.**

$$
\begin{aligned}
(\text{左辺}) &= \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots (a_{k\sigma(k)} + a'_{k\sigma(k)}) \cdots a_{N\sigma(N)} \\
&= \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{k\sigma(k)} \cdots a_{N\sigma(N)} \\
&\quad + \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a'_{k\sigma(k)} \cdots a_{N\sigma(N)} \\
&= (\text{右辺}).
\end{aligned}
$$

■

**(III)**　$j < k$ とすると

$$
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1N} \\
\vdots & \vdots & & \vdots \\
a_{j1} & a_{j2} & \cdots & a_{jN} \\
\vdots & \vdots & & \vdots \\
a_{k1} & a_{k2} & \cdots & a_{kN} \\
\vdots & \vdots & & \vdots \\
a_{N1} & a_{N2} & \cdots & a_{NN}
\end{vmatrix}
= -
\begin{vmatrix}
a_{11} & a_{12} & \cdots & a_{1N} \\
\vdots & \vdots & & \vdots \\
a_{k1} & a_{k2} & \cdots & a_{kN} \\
\vdots & \vdots & & \vdots \\
a_{j1} & a_{j2} & \cdots & a_{jN} \\
\vdots & \vdots & & \vdots \\
a_{N1} & a_{N2} & \cdots & a_{NN}
\end{vmatrix}.
$$

**証明.** $\sigma = \begin{pmatrix} 1 & \cdots & j & \cdots & k & \cdots & N \\ i_1 & \cdots & i_j & \cdots & i_k & \cdots & i_N \end{pmatrix}$ に対して,

$$
\sigma' = \begin{pmatrix} 1 & \cdots & j & \cdots & k & \cdots & N \\ i_1 & \cdots & i_k & \cdots & i_j & \cdots & i_N \end{pmatrix}
$$

とおくと,

$$
\begin{aligned}
&(i_j, i_k) \begin{pmatrix} 1 & \cdots & j & \cdots & k & \cdots & N \\ i_1 & \cdots & i_k & \cdots & i_j & \cdots & i_N \end{pmatrix} \\
&= \begin{pmatrix} 1 & \cdots & j & \cdots & k & \cdots & N \\ i_1 & \cdots & i_j & \cdots & i_k & \cdots & i_N \end{pmatrix}
\end{aligned}
$$

より $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma)$ である．また

$$
\operatorname{sgn}(\sigma')\, a_{1\sigma'(1)} \cdots a_{j\sigma'(j)} \cdots a_{k\sigma'(k)} \cdots a_{N\sigma'(N)}
$$

$$= -\mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(k)} \cdots a_{k\sigma(j)} \cdots a_{N\sigma(N)}$$
$$= -\mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{k\sigma(j)} \cdots a_{j\sigma(k)} \cdots a_{N\sigma(N)}$$

である．$S_N = \{\sigma' : \sigma \in S_N\}$ であるから，

$$\begin{aligned}
\text{(III) の左辺} &= \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{k\sigma(k)} \cdots a_{N\sigma(N)} \\
&= \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma')\, a_{1\sigma'(1)} \cdots a_{j\sigma'(j)} \cdots a_{k\sigma'(k)} \cdots a_{N\sigma'(N)} \\
&= -\sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{k\sigma(j)} \cdots a_{j\sigma(k)} \cdots a_{N\sigma(N)} \\
&= \text{(III) の右辺}.
\end{aligned}$$

■

**(IV)** 行列 $A$ の二つの行が同じであれば，すなわちある $j < k$ に対して $a_{jl} = a_{kl}\,(l = 1, \cdots, N)$ であるとき，

$$|A| = 0.$$

**証明.** (III) より $|A| = -|A|$ であるから $|A| = 0$ が成り立つ． ■

**(V)** $j < k$ とスカラー $t$ に対して

$$\begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} + ta_{k1} & \cdots & a_{jN} + ta_{kN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1}+ta_{j1} & \cdots & a_{kN}+ta_{jN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}.$$

**証明.** 最初の等式のみ示す. 他も同様.

(II), (I), (IV) より

$$\begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1}+ta_{k1} & \cdots & a_{jN}+ta_{kN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ ta_{k1} & \cdots & ta_{kN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}$$

$$
= \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} + t \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jN} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kN} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix}.
$$

■

**(VI)**　$A$ を $N$ 次正方行列としたとき，

$$
\left|A^T\right| = |A|
$$

である．

**証明.** 定理 5.8 より

$$
\begin{aligned}
\left|A^T\right| &= \begin{vmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{vmatrix} = \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(N)N} \\
&= \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{N\sigma(N)} = \begin{vmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1N} & \cdots & a_{NN} \end{vmatrix} \\
&= |A|.
\end{aligned}
$$

■

これらの公式を用いると，行列式の計算が簡単になることがある．例題 5.1(1) をここで述べた公式を用いて計算してみよう．なお，以下では，たとえば

$$
|A| \underset{R_j + \alpha R_k,\ R_l + \beta R_m}{=} |B|
$$

により，等号の左側にある行列式の第 $j$ 行に第 $k$ 行の $\alpha$ 倍を加え，引き続き第 $l$ 行に第 $m$ 行の $\beta$ 倍を加えると等号の右側の行列式になることを表すものとする．

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & 1 & 3 \\ 3 & 2 & 4 \end{vmatrix} \underset{R_2-2R_1,\ R_3-3R_1}{=} \begin{vmatrix} 1 & 3 & -2 \\ 0 & -5 & 7 \\ 0 & -7 & 10 \end{vmatrix}$$

$$\underset{R_3-\frac{7}{5}R_2}{=} \begin{vmatrix} 1 & 3 & -2 \\ 0 & -5 & 7 \\ 0 & 0 & \frac{1}{5} \end{vmatrix} \underset{\text{例題 } 5.2(2)}{=} 1 \times (-5) \times \frac{1}{5} = -1.$$

**問題 5.3** $\begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix}$ を求めよ．

### 5.4.2　列に関する変形

(I) ～ (V) は行の変形に関する公式であったが，列の変形についても次のような公式が成り立つ．

**(I)′**

$$\begin{vmatrix} a_{11} & \cdots & ta_{1k} & \cdots & a_{1N} \\ a_{21} & \cdots & ta_{2k} & \cdots & a_{2N} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{N1} & \cdots & ta_{Nk} & \cdots & a_{NN} \end{vmatrix} = t \begin{vmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2k} & \cdots & a_{2N} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{N1} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix}.$$

**証明.** 定理 5.8 より

$$\begin{aligned}(\text{左辺}) &= \sum_{\sigma \in S_N} a_{\sigma(1)1} \cdots (ta_{\sigma(k)k}) \cdots a_{\sigma(N)N} \\ &= t \sum_{\sigma \in S_N} a_{\sigma(1)1} \cdots a_{\sigma(k)k} \cdots a_{\sigma(N)N} = (\text{右辺}).\end{aligned}$$

■

以下の公式も，定理 5.8 を用いれば (II) ～ (V) と同様に証明することができる．

**(II)′**

$$\begin{vmatrix} a_{11} & \cdots & a_{1k}+a'_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nk}+a'_{Nk} & \cdots & a_{NN} \end{vmatrix}$$
$$= \begin{vmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a'_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a'_{Nk} & \cdots & a_{NN} \end{vmatrix}.$$

**(III)′** $j < k$ とすると

$$\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix}$$
$$= - \begin{vmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1j} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nk} & \cdots & a_{Nj} & \cdots & a_{NN} \end{vmatrix}.$$

**(IV)′** 行列 $A$ の二つの列が同じであれば, すなわちある $j < k$ に対して $a_{lj} = a_{lk}$, $l = 1, \cdots, N$ であるとき,

$$|A| = 0.$$

**(V)′** $j < k$ とスカラー $t$ に対して

$$\begin{vmatrix} a_{11} & \cdots & a_{1j}+ta_{1k} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj}+ta_{Nk} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix}$$
$$= \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k}+ta_{1j} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{Nk}+ta_{Nj} & \cdots & a_{NN} \end{vmatrix}$$

$$
= \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix}.
$$

**例題 5.3**

$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{2N} & \cdots & a_{NN} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2N} \\ \vdots & \ddots & \vdots \\ a_{2N} & \cdots & a_{NN} \end{vmatrix}.
$$

**答.** 列の入れ替えに関する公式 (III)$'$ を繰り返し用いて

$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ 0 & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{2N} & \cdots & a_{NN} \end{vmatrix} = (-1)^{N-1} \begin{vmatrix} a_{12} & \cdots & a_{1N} & a_{11} \\ a_{22} & \cdots & a_{2N} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{2N} & \cdots & a_{NN} & 0 \end{vmatrix}.
$$

次に行の入れ替えに関する公式 (III) を繰り返し用い，さらに例題 5.2(3) を適用して

$$
\begin{vmatrix} a_{12} & \cdots & a_{1N} & a_{11} \\ a_{22} & \cdots & a_{2N} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{2N} & \cdots & a_{NN} & 0 \end{vmatrix} = (-1)^{N-1} \begin{vmatrix} a_{22} & \cdots & a_{2N} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{2N} & \cdots & a_{NN} & 0 \\ a_{12} & \cdots & a_{1N} & a_{11} \end{vmatrix}
$$

$$
= (-1)^{N-1} a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2N} \\ \vdots & \ddots & \vdots \\ a_{2N} & \cdots & a_{NN} \end{vmatrix}
$$

を得る．よって例題が示された． ■

### 5.4.3　行列式のベクトルによる表示

行列

$$
A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \tag{5.9}
$$

に対して

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{N1} \end{pmatrix}, \quad \cdots, \quad \mathbf{a}_N = \begin{pmatrix} a_{1N} \\ \vdots \\ a_{NN} \end{pmatrix}$$

とし，このベクトルを用いて，行列 $A$ を

$$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N) \tag{5.10}$$

と表す．これを行列のベクトルによる表示という．そして行列 (5.9) の行列式を

$$\det(\mathbf{a}_1, \cdots, \mathbf{a}_N) \quad \text{あるいは} \quad D(\mathbf{a}_1, \cdots, \mathbf{a}_N)$$

と表す．このように記しておくと，しばしば行列式の計算をする際に記述が簡明になる．たとえば (I)′ ～ (V)′ は次のように記される．

**(I)′**

$$D(\mathbf{a}_1, \cdots, t\mathbf{a}_k, \cdots, \mathbf{a}_N) = tD(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N).$$

**(II)′**

$$\begin{aligned} &D(\mathbf{a}_1, \cdots, \mathbf{a}_k + \mathbf{a}'_k, \cdots, \mathbf{a}_N) \\ &\quad = D(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N) + D(\mathbf{a}_1, \cdots, \mathbf{a}'_k, \cdots, \mathbf{a}_N). \end{aligned}$$

**(III)′**　$j < k$ とすると

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N) = -D(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_N).$$

**(IV)′**　ある $j < k$ に対して $\mathbf{a}_j = \mathbf{a}_k$ であれば,

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N) = 0.$$

**(V)′**　$j < k$ とスカラー $t$ に対して

$$\begin{aligned} D(\mathbf{a}_1, \cdots, \mathbf{a}_j + t\mathbf{a}_k, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N) &= D(\mathbf{a}_1, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N), \\ D(\mathbf{a}_1, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_k + t\mathbf{a}_j, \cdots, \mathbf{a}_N) &= D(\mathbf{a}_1, \cdots, \mathbf{a}_j, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N). \end{aligned}$$

ここで行列の積の行列式に関する基本的な公式を証明しておく．

**定理 5.9**　$A$, $B$ を $N$ 次正方行列とすると

$$|AB| = |A|\,|B|$$

である．

**証明．**

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}=(\mathbf{a}_1,\cdots,\mathbf{a}_N),$$

$$B=\begin{pmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{pmatrix}$$

とする．このとき

$$\begin{aligned} AB &= \begin{pmatrix} \sum_{n=1}^{N} a_{1n}b_{n1} & \cdots & \sum_{n=1}^{N} a_{1n}b_{nN} \\ \vdots & \ddots & \vdots \\ \sum_{n=1}^{N} a_{Nn}b_{n1} & \cdots & \sum_{n=1}^{N} a_{Nn}b_{nN} \end{pmatrix} \\ &= \left(\sum_{n=1}^{N} b_{n1}\mathbf{a}_n, \cdots, \sum_{n=1}^{N} b_{nN}\mathbf{a}_n\right) \end{aligned}$$

と表される．したがって，

$$\begin{aligned} |AB| &= D\left(\sum_{n=1}^{N} b_{n1}\mathbf{a}_n, \cdots, \sum_{n=1}^{N} b_{nN}\mathbf{a}_n\right) \\ &= \sum_{n_1=1}^{N} \cdots \sum_{n_N=1}^{N} b_{n_1 1}\cdots b_{n_N N} D(\mathbf{a}_{n_1},\cdots,\mathbf{a}_{n_N}) \\ &= \sum_{n_1,\cdots,n_N=1,\, n_i\neq n_j,\, i\neq j}^{N} b_{n_1 1}\cdots b_{n_N N} D(\mathbf{a}_{n_1},\cdots,\mathbf{a}_{n_N}) \\ &= \sum_{n_1,\cdots,n_N=1,\, n_i\neq n_j,\, i\neq j}^{N} b_{n_1 1}\cdots b_{n_N N}\, \mathrm{sgn}\begin{pmatrix} 1 & \cdots & N \\ n_1 & \cdots & n_N \end{pmatrix} D(\mathbf{a}_1,\cdots,\mathbf{a}_N) \\ &= |B|\,|A| = |A|\,|B|. \end{aligned}$$

■

次の系もよく使われる公式である．

**系 5.10**　$N$ 次正方行列 $A$ が正則ならば $|A| \neq 0$ であり，次の等式が成り立つ．

$$\left|A^{-1}\right| = |A|^{-1}.$$

**証明.** $I_N = A^{-1}A$ より，$1 = |A^{-1}|\,|A|$．■

**系 5.11**　$A$ を $N$ 次正方行列で，$P$ を正則な $N$ 次正方行列とする．このとき

$$\left|P^{-1}AP\right| = |A|.$$

**証明.** $|P^{-1}AP| = |P^{-1}|\,|A|\,|P| = |P|^{-1}\,|A|\,|P| = |A|$．■

**問題 5.4**　$A \in \boldsymbol{K}^{N\times N}$ を正則行列であるとする．基本行列 $E_1, \cdots, E_n$ により

$$E_n \cdots E_1 A = \mathrm{Trg}^{+}_{N,N}(\alpha_1, \cdots, \alpha_N), \quad \alpha_1 \cdots \alpha_N \neq 0$$

と上三角行列に変形できる．

# 第6章　行列式の余因子展開とその応用

行列式の計算をする際に，行列式の余因子展開と呼ばれる展開を使うと便利なことが多い．この章では行列式の余因子展開を述べ，さらにその応用として逆行列の行列式を用いた表し方，連立1次方程式の解を求めるクラメルの公式を証明する．

**ポイント6**　余因子展開は行列式の計算をする際の便利な公式である．使えるようにしておくことが望ましい．

## 6.1　余因子展開

はじめに $(3,3)$ 行列の行列式の場合に展開公式を導いておく．前節で証明した公式を用いて次のような変形ができる．

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}+0+0 & 0+a_{12}+0 & 0+0+a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{12} & 0 & 0 \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{13} & 0 & 0 \\ a_{23} & a_{21} & a_{22} \\ a_{33} & a_{31} & a_{32} \end{vmatrix}.$$

さらに，例題5.3より

$$\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

である．他の部分についても同様のことが成り立つので，次を得る．

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{6.1}$$

これにより3次正方行列の行列式が2次正方行列の行列式の計算に帰着された．上の公式が余因子展開と呼ばれるものである．これは次のように $N$ 次正方行列に一般化される．

$N$ 次正方行列 $A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$ のうち第 $i$ 行と第 $j$ 列を取り除いてできる $N-1$ 次正方行列を $A^{(ij)}$ と表す．すなわち

$$A^{(ij)} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,N} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{N,j-1} & a_{N,j+1} & \cdots & a_{NN} \end{pmatrix}$$

である．この行列の行列式に符号 $(-1)^{i+j}$ をかけたもの

$$\widetilde{A}_{ij} = (-1)^{i+j} \left| A^{(ij)} \right|$$

を $A$ の**第$(i,j)$ 余因子**という．

余因子を用いると (6.1) は次のように表すことができる．

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \left| A^{(11)} \right| - a_{12} \left| A^{(12)} \right| + a_{13} \left| A^{(13)} \right| \\ &= a_{11}\widetilde{A}_{11} + a_{12}\widetilde{A}_{12} + a_{13}\widetilde{A}_{13}. \end{aligned}$$

一般に次のことが成り立つ．

**定理 6.1**　$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$ とする．$1 \leq i \leq N$ に対して

$$|A| = \sum_{k=1}^{N} (-1)^{i+k} a_{ik} \left| A^{(ik)} \right| = \sum_{k=1}^{N} a_{ik} \widetilde{A}_{ik} \tag{6.2}$$

が成り立つ．また，$1 \leq j \leq N$ に対して

$$|A| = \sum_{k=1}^{N} (-1)^{k+j} a_{kj} \left| A^{(kj)} \right| = \sum_{k=1}^{N} a_{kj} \widetilde{A}_{kj} \tag{6.3}$$

が成り立つ．

(6.2) を $|A|$ の**第 $i$ 行に関する余因子展開**，また (6.3) を**第 $j$ 列に関する余因子展開**という．

**証明.** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とおく．$\mathbf{a}_j = \sum_{i=1}^{N} a_{ij} \mathbf{e}_i$ であるから，

$$\begin{aligned}
|A| &= D(\mathbf{a}_1, \cdots, \mathbf{a}_N) = D(\mathbf{a}_1, \cdots, \sum_{i=1}^{N} a_{ij} \mathbf{e}_i, \cdots, \mathbf{a}_N) \\
&= \sum_{i=1}^{N} a_{ij} D(\mathbf{a}_1, \cdots, \overset{j\,\text{番目}}{\mathbf{e}_i}, \cdots, \mathbf{a}_N) \\
&= \sum_{i=1}^{N} (-1)^{j-1} a_{ij} D(\mathbf{e}_i, \mathbf{a}_1, \cdots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \cdots, \mathbf{a}_N) \\
&= \sum_{i=1}^{N} (-1)^{j-1} a_{ij} \begin{vmatrix} 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1N} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{iN} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{N1} & \cdots & a_{N,j-1} & a_{N,j+1} & \cdots & a_{NN} \end{vmatrix}
\end{aligned}$$

$$
= \sum_{i=1}^{N} (-1)^{j-1+i-1} a_{ij} \begin{vmatrix} 1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{iN} \\ 0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1N} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,N} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,N} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{N1} & \cdots & a_{N,j-1} & a_{N,j+1} & \cdots & a_{NN} \end{vmatrix}
$$

$$
= \sum_{i=1}^{N} (-1)^{i+j} a_{ij} \left| A^{(ij)} \right| .
$$

列に関する展開も同様にして証明できる． ■

**系 6.2**　$1 \leq k, l \leq N$ に対して

$$
\sum_{i=1}^{N} a_{ik} \widetilde{A}_{il} = \sum_{j=1}^{N} a_{kj} \widetilde{A}_{lj} = \begin{cases} 0 & (k \neq l) \\ |A| & (k = l) \end{cases} .
$$

**証明．** $k = l$ の場合は定理 6.1 より明らか．$k \neq l$ の場合を示す．$A = (a_{ij}) = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とおく．$D(\mathbf{a}_1, \cdots, \mathbf{a}_N)$ の $\mathbf{a}_l$ の位置に $\mathbf{a}_k$ を代入すると

$$
\begin{aligned}
0 &= D(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_N) \\
&= \begin{vmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1k} & \cdots & a_{1N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N1} & \cdots & a_{Nk} & \cdots & a_{Nk} & \cdots & a_{NN} \end{vmatrix} .
\end{aligned}
$$

これを第 $l$ 列に関する余因子展開をして

$$
0 = \sum_{i=1}^{N} a_{ik} \widetilde{A}_{il}
$$

を得る．同様にして，第 $k$ 行と第 $l$ 行が同じ行列を第 $l$ 行に関して余因子展開すれば残りの主張も証明できる． ■

## 6.2　余因子展開を用いた行列式の計算

行あるいは列に関する余因子展開を用いると，$(N, N)$ 行列の行列式の計算が $(N-1, N-1)$ 行列に帰着される．必要ならさらに余因子展開を繰り返し用いて，

$(N-k, N-k)$ 行列の行列式の計算に帰着することもできる．このように行列式の計算をより低次の行列の行列式に帰着させることができる．ここではこの方法による行列式の計算例を示す．

**例 6.1**　第 2 行に関する余因子展開をして

$$\begin{vmatrix} 1 & 3 & 1 \\ 3 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = -3\begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} + 0\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} - \begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix}$$
$$= -3(-1) - (-2) = 5.$$

**例 6.2**　第 3 行に関する余因子展開をして

$$\begin{vmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{vmatrix} - 0\begin{vmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{vmatrix}$$
$$+ \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} + 0\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 0 \end{vmatrix}$$
$$= \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}.$$

次に，得られた $(3,3)$ 行列の行列式の第 1 行に関する余因子展開をすれば

$$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 2\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 4,$$

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = -2.$$

したがって求める行列式の値は 2 である．

## 6.3 行列式と余因子を用いた逆行列の計算

本節では逆行列が行列式と余因子を用いて次のように表せることを示す.

**定理 6.3** $A$ を $N$ 次正方行列とする. $A$ が正則であるための必要十分条件は $|A| \neq 0$ である.

$|A| \neq 0$ のとき,

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} \widetilde{A}_{11} & \widetilde{A}_{21} & \cdots & \widetilde{A}_{N1} \\ \widetilde{A}_{12} & \widetilde{A}_{22} & \cdots & \widetilde{A}_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{A}_{1N} & \widetilde{A}_{2N} & \cdots & \widetilde{A}_{NN} \end{pmatrix} \tag{6.4}$$

が成り立つ (**注意: 右辺において** $\widetilde{A}_{ij}$ **の配列は** $i$ **が列,** $j$ **が行を表す配置になっていることに注意せよ**).

**証明.** 系 5.10 より $A$ が正則であれば $|A| \neq 0$ である. 逆を示す. $|A| \neq 0$ とする. (6.4) の右辺を $B$ とおくと, 定理 6.1 と系 6.2 より

$$\begin{aligned} BA &= \frac{1}{|A|} \begin{pmatrix} \widetilde{A}_{11} & \cdots & \widetilde{A}_{N1} \\ \vdots & \ddots & \vdots \\ \widetilde{A}_{1N} & \cdots & \widetilde{A}_{NN} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \\ &= \left( \frac{1}{|A|} \sum_{i=1}^{N} a_{il} \widetilde{A}_{ik} \right)_{k,l} = I_N \end{aligned}$$

を得る. また

$$\begin{aligned} AB &= \frac{1}{|A|} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \begin{pmatrix} \widetilde{A}_{11} & \cdots & \widetilde{A}_{N1} \\ \vdots & \ddots & \vdots \\ \widetilde{A}_{1N} & \cdots & \widetilde{A}_{NN} \end{pmatrix} \\ &= \left( \sum_{j=1}^{N} a_{kj} \frac{\widetilde{A}_{lj}}{|A|} \right)_{k,l} = I_N \end{aligned}$$

も得られる. よって $B = A^{-1}$ である. ■

この定理から次のことがわかる.

**系 6.4** $A$ を $N$ 次正方行列とする. もしある $N$ 次正方行列 $B$ で $AB = I_N$ と

なるものが存在するならば，$A$ は正則であり，$B=A^{-1}$ である．また，もしある $N$ 次正方行列 $B$ で $BA=I_N$ となるものが存在するならば，$A$ は正則であり，$B=A^{-1}$ である．

**証明.** 前半の主張を示す．$1=|AB|=|A|\,|B|$ より $|A|\neq 0$ であるから，$A$ は正則である．さらに

$$B=\left(A^{-1}A\right)B=A^{-1}(AB)=A^{-1}I_N=A^{-1}$$

となっている．後半の主張も同様にして証明できる．■

**例題 6.1**　(6.4) を用いて $A=\begin{pmatrix}2&3&1\\4&1&3\\3&1&2\end{pmatrix}$ の逆行列を求めよ．

**答.**

$$|A|=\begin{vmatrix}2&3&1\\4&1&3\\3&1&2\end{vmatrix}=\begin{vmatrix}1&3&1\\1&1&3\\1&1&2\end{vmatrix}=\begin{vmatrix}1&2&1\\1&0&3\\1&0&2\end{vmatrix}$$

$$=-2\begin{vmatrix}1&3\\1&2\end{vmatrix}=-2(2-3)=2.$$

また

$$\widetilde{A}_{11}=\begin{vmatrix}1&3\\1&2\end{vmatrix}=-1,\quad \widetilde{A}_{12}=-\begin{vmatrix}4&3\\3&2\end{vmatrix}=1,\quad \widetilde{A}_{13}=\begin{vmatrix}4&1\\3&1\end{vmatrix}=1,$$

$$\widetilde{A}_{21}=-\begin{vmatrix}3&1\\1&2\end{vmatrix}=-5,\quad \widetilde{A}_{22}=\begin{vmatrix}2&1\\3&2\end{vmatrix}=1,\quad \widetilde{A}_{23}=-\begin{vmatrix}2&3\\3&1\end{vmatrix}=7,$$

$$\widetilde{A}_{31}=\begin{vmatrix}3&1\\1&3\end{vmatrix}=8,\quad \widetilde{A}_{32}=-\begin{vmatrix}2&1\\4&3\end{vmatrix}=-2,\quad \widetilde{A}_{33}=\begin{vmatrix}2&3\\4&1\end{vmatrix}=-10.$$

したがって，

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix}\widetilde{A}_{11}&\widetilde{A}_{21}&\widetilde{A}_{31}\\\widetilde{A}_{12}&\widetilde{A}_{22}&\widetilde{A}_{32}\\\widetilde{A}_{13}&\widetilde{A}_{23}&\widetilde{A}_{33}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}-1&-5&8\\1&1&-2\\1&7&-10\end{pmatrix}.$$

■

## 6.4 行列式と連立 1 次方程式の解法

連立 1 次方程式

$$\begin{cases} a_{11}x_1 + \cdots + a_{1N}x_N = b_1 \\ a_{21}x_1 + \cdots + a_{2N}x_N = b_2 \\ \qquad \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N = b_M \end{cases} \tag{6.5}$$

で，特に $b_1 = \cdots = b_M = 0$ であるもの，すなわち

$$\begin{cases} a_{11}x_1 + \cdots + a_{1N}x_N = 0 \\ a_{21}x_1 + \cdots + a_{2N}x_N = 0 \\ \qquad \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N = 0 \end{cases} \tag{6.6}$$

を**斉次連立 1 次方程式**という．これに対して，(6.5) で $(b_1, \cdots, b_M) \neq (0, \cdots, 0)$ の場合を**非斉次連立 1 次方程式**という．斉次連立 1 次方程式 (6.6) は必ず $x_1 = \cdots = x_N = 0$ を一つの解としてもつ．この解を方程式 (6.6) の**自明解**あるいは**自明な解**という．自明解でない解を**非自明解**あるいは**非自明な解**という．

斉次連立 1 次方程式 (6.6) は非自明な解をもつ場合ともたない場合がある．これについて，$M = N$ の場合は係数行列

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$$

の行列式を用いた判定法がある．

**定理 6.5** $M = N$ の場合，斉次連立 1 次方程式 (6.6) が非自明な解をもつための必要十分条件は $|A| = 0$ となることである．

**証明.** もし $|A| \neq 0$ ならば $A$ は逆行列をもつ．したがって，$A\mathbf{x} = \mathbf{0}$ をみたす $\mathbf{x}$ は必ず

$$\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$$

であるから, 非自明な解を持たない．したがって, 非自明な解を持つならば, $|A| = 0$ である．

逆を示す．$|A| = 0$ とする．$N$ に関する帰納法を使う．$N = 1$ の場合, $|A| = a_{11}$

であるから，任意の $x_1 \neq 0$ が $a_{11}x_1 = 0$ の解になっている．したがって非自明解をもつ．$N = n-1$ の場合に成立したとして，$N = n$ の場合を示す．すべての $a_{ij}$ が 0 ならば，明らかにすべての $\mathbf{x} \in \boldsymbol{K}^n$ に対して $A\mathbf{x} = 0$ である．したがって $a_{ij} \neq 0$ なる $a_{ij}$ が存在する場合に証明すればよい．必要なら，行ならびに列の入れ替えにより，$a_{ij}$ を $(1,1)$ 成分の位置にもってくることができる．したがって，$a_{11} \neq 0$ の場合を扱えばよいことがわかる．このとき，(6.6) の第 1 式より

$$x_1 = -\frac{1}{a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n) \tag{6.7}$$

である．これを第 $i$ 式 $a_{i1}x_1 + \cdots + a_{in}x_n = 0\,(2 \leq i \leq n)$ に代入すると

$$-\frac{a_{i1}}{a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n) + a_{i2}x_2 + \cdots + a_{in}x_n = 0$$

であるから，

$$\left(a_{i2} - \frac{a_{i1}a_{12}}{a_{11}}\right)x_2 + \cdots + \left(a_{in} - \frac{a_{i1}a_{1n}}{a_{11}}\right)x_n = 0 \quad (i = 2, \cdots, n)$$

である．したがって，

$$B = \begin{pmatrix} a_{22} - \dfrac{a_{21}a_{12}}{a_{11}} & \cdots & a_{2n} - \dfrac{a_{21}a_{1n}}{a_{11}} \\ \vdots & \ddots & \vdots \\ a_{n2} - \dfrac{a_{n1}a_{12}}{a_{11}} & \cdots & a_{nn} - \dfrac{a_{n1}a_{1n}}{a_{11}} \end{pmatrix}, \quad \mathbf{x}' = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}$$

とすると $B\mathbf{x}' = \mathbf{0}_{n-1}$ である．ところで

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}}a_{12} & \cdots & a_{2n} - \dfrac{a_{21}}{a_{11}}a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} - \dfrac{a_{n1}}{a_{11}}a_{12} & \cdots & a_{nn} - \dfrac{a_{n1}}{a_{11}}a_{1n} \end{vmatrix}$$

$$\begin{pmatrix} \text{第 2 行} - \dfrac{a_{21}}{a_{11}} \times \text{第 1 行} \\ \vdots \\ \text{第 } n \text{ 行} - \dfrac{a_{n1}}{a_{11}} \times \text{第 1 行} \end{pmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{vmatrix} = a_{11}\,|B|$$

である．仮定より $|A| = 0, a_{11} \neq 0$ であるから $|B| = 0$ となる．したがって帰納法の仮定から方程式 $B\mathbf{y} = \mathbf{0}_{n-1}$ は非自明解 $\mathbf{x}'$ をもつ．これを用いて (6.7) により $x_1$ を定めれば，$\mathbf{x} = (x_j) \in \boldsymbol{K}^n$ が方程式 (6.6) の非自明解であることがわかる． ■

連立 1 次方程式 (6.5) は

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}$$

とおけば，$A\mathbf{x} = \mathbf{b}$ と表せる．したがって，$M = N$ の場合，もしも $|A| \neq 0$ であるならば，この方程式の解は $A$ の逆行列を用いて

$$\mathbf{x} = A^{-1}\mathbf{b}$$

によって与えられる．さらに次のことが成り立つ．

**定理 6.6（クラメルの公式）** $M = N$ とする．連立 1 次方程式 (6.5) について，$|A| \neq 0$ の場合，(6.5) の解 $x_1, \cdots, x_N$ は

$$x_i = \frac{1}{|A|} \det(\mathbf{a}_1, \cdots, \mathbf{a}_{i-1}, \overset{i\text{ 番目}}{\mathbf{b}}, \mathbf{a}_{i+1} \cdots, \mathbf{a}_N)$$

$(i = 1, \cdots, N)$ で与えられる．

**証明.** $\mathbf{x}$ が $A\mathbf{x} = \mathbf{b}$ をみたすとすると，

$$\mathbf{b} = (\mathbf{a}_1, \cdots, \mathbf{a}_N) \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \sum_{k=1}^{N} x_k \mathbf{a}_k$$

であるから，

$$\det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\mathbf{b}}, \cdots, \mathbf{a}_N) = \det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\sum_{k=1}^{N} x_k \mathbf{a}_k}, \cdots, \mathbf{a}_N)$$
$$= \sum_{k=1}^{N} x_k \det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\mathbf{a}_k}, \cdots, \mathbf{a}_N)$$

である．$k \neq i$ の場合，$k$ 番目にも $\mathbf{a}_k$ があるので，

$$\det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\mathbf{a}_k}, \cdots, \mathbf{a}_N) = 0$$

である．また $k = i$ の場合は，

$$\det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\mathbf{a}_k}, \cdots, \mathbf{a}_N) = |A|$$

となっている．したがって

$$\det(\mathbf{a}_1, \cdots, \overset{i\text{ 番目}}{\mathbf{b}}, \cdots, \mathbf{a}_N) = x_i\,|A|$$

であり，これより定理が得られる． ■

具体的な連立 1 次方程式を解く場合，あるいは逆行列を求める際，クラメルの公式や公式 (6.4) を用いて計算すると計算量が多くなる．むしろガウスの消去法などを用いた方がよい．しかしクラメルの公式，公式 (6.4) は線形代数の理論的研究には使われる．たとえば第 10.5 節の定理 10.22 や以下の問題のうち問題 6.3 などはクラメルの公式を用いれば容易に示すことができる．

**問題 6.1**

$$\begin{cases} 2x_1 + 3x_2 + \phantom{3}x_3 = 1 \\ 4x_1 + \phantom{3}x_2 + 3x_3 = 2 \\ 3x_1 + \phantom{3}x_2 + 2x_3 = 1 \end{cases} \tag{6.8}$$

を逆行列を求めることにより解け．

**問題 6.2**　クラメルの公式を用いて，連立 1 次方程式 (6.8) を解け．

**問題 6.3**　(6.5) においてすべての $a_{ij}, b_j$ が整数であるとする．$A = (a_{ij})$ が $|\det A| = 1$ をみたすならば，方程式 (6.5) は整数解をもつことを示せ．

# 第7章　いろいろな行列の行列式

この節ではいくつかの応用上重要な行列の行列式の計算を述べる.

**ポイント7**　本章ではいくつかの有用な行列の行列式について述べる. 特に特性多項式は固有値の問題を扱う際に重要な役割を果たす. また小行列式も後で使われる.

## 7.1　ファンデルモンド行列式と補間多項式への応用

$$V = V(x_1, \cdots, x_N) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_N \\ x_1^2 & x_2^2 & \cdots & x_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{N-1} & x_2^{N-1} & \cdots & x_N^{N-1} \end{pmatrix}$$

を $N$ 次**ファンデルモンド行列**という. 本節ではファンデルモンド行列の行列式を求める. またその応用例として, 多項式による補間問題について述べる. これは

**[多項式による補間問題]**

$x_1, \cdots, x_N \in \boldsymbol{K}$ を相異なる点とする. 任意に $w_1, \cdots, w_N \in \boldsymbol{K}$ が与えられたとき,

$$p(x_1) = w_1, \quad \cdots, \quad p(x_N) = w_N$$

をみたす $N-1$ 次多項式 $p(z)$ を見出せ.

という問題である.

まずファンデルモンド行列の行列式を求める．

**定理 7.1**　$N$ 次ファンデルモンド行列の行列式は

$$|V| = \prod_{i<j}(x_j - x_i) = (-1)^{\frac{N(N-1)}{2}} \prod_{i<j}(x_i - x_j)$$

である．ただし，ここで $\prod_{i<j}(x_j - x_i)$ はすべての $1 \le i < j \le N$ にわたって $(x_j - x_i)$ の積をとることを意味する記号である．$\prod_{i<j}(x_i - x_j)$ も同じように定義される．

**証明.** $N = 2$ の場合は明らかに

$$\begin{vmatrix} 1 & 1 \\ x_1 & x_2 \end{vmatrix} = x_2 - x_1.$$

$N = 3$ の場合，

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ x_1 & x_2 - x_1 & x_3 - x_1 \\ x_1^2 & x_2^2 - x_1^2 & x_3^2 - x_1^2 \end{vmatrix} = \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ x_2^2 - x_1^2 & x_3^2 - x_1^2 \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & 1 \\ x_2 + x_1 & x_3 + x_1 \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & 1 \\ x_2 & x_3 \end{vmatrix}. \tag{7.1}$$

ここで $\begin{vmatrix} 1 & 1 \\ x_2 & x_3 \end{vmatrix}$ は 1 次少ないファンデルモンド行列の行列式であることに注意しておく．したがって

$$(7.1) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2) = \prod_{i<j}(x_i - x_j)$$

であることがわかる．

$N = 3$ の場合の考察をもとに，一般の場合を帰納法で証明する．$n - 1$ の場合に正しいとして，$n$ の場合を証明する．第 $k$ 列から第 1 列を引いて，

$$|V| = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ x_1 & x_2 - x_1 & \cdots & x_n - x_1 \\ x_1^2 & x_2^2 - x_1^2 & \cdots & x_n^2 - x_1^2 \\ x_1^3 & x_2^3 - x_1^3 & \cdots & x_n^3 - x_1^3 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} - x_1^{n-1} & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix} \tag{7.2}$$

を得るが，ここで第 1 行に関して余因子展開をして

$$(7.2) = \begin{vmatrix} x_2 - x_1 & \cdots & x_n - x_1 \\ x_2^2 - x_1^2 & \cdots & x_n^2 - x_1^2 \\ x_2^3 - x_1^3 & \cdots & x_n^3 - x_1^3 \\ \vdots & \ddots & \vdots \\ x_2^{n-1} - x_1^{n-1} & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix} = (x_2 - x_1)\cdots(x_n - x_1)$$

$$\times \begin{vmatrix} 1 & \cdots & 1 \\ x_2 + x_1 & \cdots & x_n + x_1 \\ x_2^2 + x_2x_1 + x_1^2 & \cdots & x_n^2 + x_nx_1 + x_1^2 \\ \vdots & \ddots & \vdots \\ \sum_{j=0}^{n-2} x_2^{n-2-j}x_1^j & \cdots & \sum_{j=0}^{n-2} x_n^{n-2-j}x_1^j \end{vmatrix}. \tag{7.3}$$

ここで各 $k = 2, \cdots, n-1$ に対して，( 第 $k$ 行 ) − ( 第 $k-1$ 行 ) $\times x_1$ の操作を行えば，

$$\sum_{j=0}^{k-1} x_m^{k-1-j}x_1^j - \left(\sum_{j=0}^{k-2} x_m^{k-2-j}x_1^j\right)x_1$$

$$= x_m^{k-1} + \sum_{j=1}^{k-1} x_m^{k-1-j}x_1^j - \sum_{j=0}^{k-2} x_m^{k-2-j}x_1^{j+1} = x_m^{k-1}$$

であるから次を得る.

$$(7.3) = (x_2 - x_1)\cdots(x_n - x_1)\begin{vmatrix} 1 & \cdots & 1 \\ x_2 & \cdots & x_n \\ x_2^2 & \cdots & x_3^2 \\ \vdots & \ddots & \vdots \\ x_2^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}. \tag{7.4}$$

したがって帰納法の仮定から，

$$\begin{aligned}(7.4) &= (x_2 - x_1)\cdots(x_n - x_1)\prod_{\substack{i<j \\ i,j\geq 2}}(x_j - x_i) \\ &= \prod_{i<j}(x_j - x_i).\end{aligned}$$

■

次に補間問題を考えよう．

$x_1, \cdots, x_N \in \boldsymbol{K}$ を相異なる点とする．任意に $w_1, \cdots, w_N \in \boldsymbol{K}$ が与えられたとき，$p(x_j) = w_j$, $j = 1, \cdots, N$ をみたす $N-1$ 次多項式 $p(z)$ を補間多項式という．補間多項式を構成するためには，多項式

$$p(z) = a_1 + a_2 z + \cdots + a_N z^{N-1}$$

の係数 $a_1, \cdots, a_N$ を $p(x_j) = w_j$ となるように定めればよい．そのためには，連立 1 次方程式

$$\begin{gathered} a_1 + x_1 a_2 + \cdots + x_1^{N-1} a_N = p(x_1) = w_1 \\ \vdots \\ a_1 + x_N a_2 + \cdots + x_N^{N-1} a_N = p(x_N) = w_N \end{gathered}$$

の解 $a_1, \cdots, a_N$ を見出せばよい．この連立 1 次方程式はファンデルモンド行列を用いて

$$V(x_1, \cdots, x_N)^T \mathbf{a} = \mathbf{w}$$

と表せる．いま $x_1, \cdots, x_N$ は相異なる点であるから，ファンデルモンドの行列式は

$$\left|V(x_1, \cdots, x_N)^T\right| = |V(x_1, \cdots, x_N)| = \prod_{i<j}(x_i - x_i) \neq 0.$$

したがって $V(x_1, \cdots, x_N)^T$ は正則であり，求める係数は

$$\mathbf{a} = \left(V\left(x_1, \cdots, x_N\right)^T\right)^{-1}\mathbf{w}$$

により与えられる．この係数を具体的に求めてみよう．

$\Pi_i(t) = \displaystyle\prod_{j=1, j\neq i}^{N} (t - x_j)$ とおく．これは $t$ に関する $N-1$ 次の多項式である．$t$ について展開したものを

$$\Pi_i(t) = \sum_{j=1}^{N} c_{ij} t^{j-1} \tag{7.5}$$

とおく．明らかに定義より $\Pi_i(x_k) = 0,\ k \neq i$, かつ $\Pi_i(x_i) \neq 0$ である．

$$c'_{ij} = \frac{c_{ij}}{\Pi_i(x_i)}$$

とおくと

$$\sum_{j=1}^{N} c'_{ij} x_k^{j-1} = \sum_{j=1}^{N} \frac{c_{ij}}{\Pi_i(x_i)} x_k^{j-1} = \frac{\Pi_i(x_k)}{\Pi_i(x_i)} = \delta_{ik}$$

であるから，$V = V(x_1, \cdots, x_N)$ とおくと

$$V^{-1} = (c'_{ij})$$

が得られる．ゆえに

$$\mathbf{a} = \left(V^T\right)^{-1}\mathbf{w} = \left(V^{-1}\right)^T\mathbf{w} = \begin{pmatrix} c'_{11} & \cdots & c'_{N1} \\ \vdots & \ddots & \vdots \\ c'_{1N} & \cdots & c'_{NN} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{i=1}^{N} c'_{i1} w_i \\ \vdots \\ \displaystyle\sum_{i=1}^{N} c'_{iN} w_i \end{pmatrix}.$$

よって求める多項式は

$$p(t) = \sum_{j=1}^{N} a_j t^{j-1} = \sum_{j=1}^{N}\sum_{i=1}^{N} c'_{ij} w_i t^{j-1} = \sum_{i=1}^{N} w_i \frac{1}{\Pi_i(x_i)} \sum_{j=1}^{N} c_{ij} t^{j-1}$$

$$= \sum_{i=1}^{N} w_i \frac{\Pi_i(t)}{\Pi_i(x_i)}$$

である．これは $N-1$ 次の**ラグランジュ補間多項式**と呼ばれているものである．

**注意 7.1**　展開式 (7.5) の係数はいわゆる基本対称式で表されている．$s_{i,j}$ を $x_1, \cdots, x_N$ から $x_i$ を除いた $N-1$ 個の変数の $j$ 次の基本対称式とすると

$$\Pi_i(t) = \sum_{j=1}^{N} (-1)^{N-j} s_{i,N-j} t^{j-1}$$

である．

**例題 7.1**　3 次の多項式 $p(x)$ で $p(1)=3,\ p(2)=2,\ p(3)=5,\ p(4)=4$ をみたすものを求めよ．

**答．** ラグランジュの補間多項式をこの場合に計算すると

$$\begin{aligned} p(x) &= 3\frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} + 2\frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)} \\ &\quad + 5\frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)} + 4\frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)} \\ &= -\frac{4}{3}x^3 + 10x^2 - \frac{65}{3}x + 16 \end{aligned}$$

である．(図 7.1 参照) ■

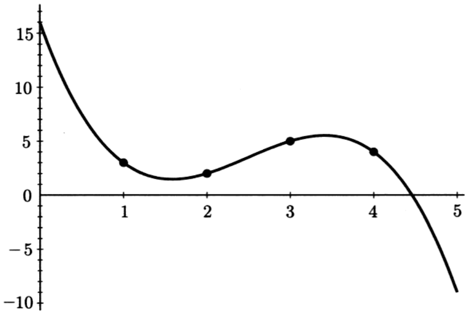



図 7.1

## 7.2　置換行列

置換行列は行列の行や列を置換する操作を実行するものである．本節では置換行列の定義とその行列式について述べる．

$$\delta_{jk}^{\mu\nu} = \begin{cases} 1 & j=\mu \quad かつ \quad k=\nu \\ 0 & その他 \end{cases}$$

とおき

$$E_{\mu\nu}^{(N)} = \left(\delta_{jk}^{\mu\nu}\right)_{\substack{j=1,\cdots,N \\ k=1,\cdots,N}}$$

とおく．これは $(\mu,\nu)$ 成分が 1 で，それ以外の成分はすべて 0 であるような $N$ 次正方行列である．すなわち

$$E_{\mu\nu}^{(N)} = \mathbf{e}_\mu^{(N)} \left(\mathbf{e}_\nu^{(N)}\right)^T$$

である．このとき $A = (a_{jk}) \in \boldsymbol{K}^{N\times M}$ に対して

$$E_{\mu\nu}^{(N)} A = \left(\sum_{n=1}^{N} \delta_{jn}^{\mu\nu} a_{nk}\right) = \left(\delta_{j\nu}^{\mu\nu} a_{\nu k}\right) = \begin{pmatrix} & O & \\ a_{\nu 1} & \cdots & a_{\nu N} \\ & O & \end{pmatrix} \tag{7.6}$$

となっているが，ただし (7.6) の最右辺では，$a_{\nu 1}, \cdots, a_{\nu N}$ は第 $\mu$ 行に並んでいる．また $E_{\mu\nu}^{(N)T} = E_{\nu\mu}^{(N)}$ であり，

$$AE_{\nu\mu}^{(N)} = \left(\sum_{n=1}^{N} a_{jn} \delta_{nk}^{\nu\mu}\right) = \left(a_{j\nu} \delta_{\nu k}^{\nu\mu}\right) = \begin{pmatrix} & a_{1\nu} & \\ O & \vdots & O \\ & a_{N\nu} & \end{pmatrix} \tag{7.7}$$

であるが，(7.7) の最右辺において $a_{1\nu}, \cdots, a_{N\nu}$ は第 $\mu$ 列に並んでいる．

$N$ 次置換 $\sigma = \begin{pmatrix} 1 & \cdots & N \\ \sigma(1) & \cdots & \sigma(N) \end{pmatrix}$ に対して，

$$P_\sigma = E_{1\sigma(1)}^{(N)} + \cdots + E_{N\sigma(N)}^{(N)}$$

とし，これを $\sigma$ **に対する置換行列**という．(7.6) より $A = (a_{jk}) \in \boldsymbol{K}^{N\times M}$ に対して

$$P_\sigma A = \begin{pmatrix} a_{\sigma(1),1} & \cdots & a_{\sigma(1),M} \\ a_{\sigma(2),1} & \cdots & a_{\sigma(2),M} \\ \vdots & \ddots & \vdots \\ a_{\sigma(N),1} & \cdots & a_{\sigma(N),M} \end{pmatrix}$$

となることがわかる．すなわち $P_\sigma A$ により $A$ の行の置換が行われる．

**例 7.1** $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ に対する置換行列は

$$P_\sigma = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

である．実際，$A = (a_{ij}) \in \boldsymbol{K}^{4\times 4}$ に対して

$$P_\sigma A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

$$= \begin{pmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix}$$

となっている．

ところで，置換行列の転置は

$$P_\sigma^T = E^{(N)}_{\sigma(1),1} + \cdots + E^{(N)}_{\sigma(N),N}$$

であるから，$A = (a_{jk}) \in K^{M\times N}$ に対して

$$AP_\sigma^T = \begin{pmatrix} a_{1,\sigma(1)} & a_{1,\sigma(2)} & \cdots & a_{1,\sigma(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,\sigma(1)} & a_{M,\sigma(2)} & \cdots & a_{M,\sigma(N)} \end{pmatrix}$$

が成り立っている．すなわち $AP_\sigma^T$ により $A$ の列の置換がなされている．

**例 7.2**　$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ とすると，

$$P_\sigma^T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

である．また

$$AP_\sigma^T = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} a_{13} & a_{11} & a_{14} & a_{12} \\ a_{23} & a_{21} & a_{24} & a_{22} \\ a_{33} & a_{31} & a_{34} & a_{32} \\ a_{43} & a_{41} & a_{44} & a_{42} \end{pmatrix}$$

が得られる．

置換行列に対して次のことが成り立つ．

**定理 7.2** $N$ 次置換 $\sigma$ に対する置換行列を $P_\sigma$ とする．このとき次のことが成り立つ．

$$P_\sigma^{-1} = P_\sigma^T,$$
$$|P_\sigma| = \left|P_\sigma^T\right| = \mathrm{sgn}(\sigma).$$

**証明.** 定義から，$P_\sigma = (p_{jk})$ とすると $p_{j\sigma(j)} = 1\,(j = 1, \cdots, N)$ であり，それ以外の成分は 0 となっている．したがって

$$|P_\sigma| = \sum_{\tau \in S_N} \mathrm{sgn}(\tau) p_{1\tau(1)} p_{2\tau(2)} \cdots p_{N\tau(N)} = \mathrm{sgn}(\sigma)$$

である．$P_\sigma^T = (q_{jk})$ とすると $q_{\sigma(j)j} = 1\,(j = 1, \cdots, N)$ であり，それ以外の成分は 0 となっている．したがって

$$\sum_{\mu=1}^{N} p_{j\mu} q_{\mu k} = p_{j\sigma(j)} q_{\sigma(j)k} = q_{\sigma(j)k} = \delta_{jk}$$

であるから，

$$P_\sigma P_\sigma^T = \left( \sum_{\mu=1}^{N} p_{j\mu} q_{\mu k} \right) = (\delta_{jk}) = I_N$$

となっている．ゆえに $P_\sigma^T = P_\sigma^{-1}$ である． ■

## 7.3　巡回行列の行列式

$$C = \mathrm{circ}\,(x_0, \cdots, x_{N-1}) = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{N-1} \\ x_{N-1} & x_0 & x_1 & \cdots & x_{N-2} \\ x_{N-2} & x_{N-1} & x_0 & \cdots & x_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_0 \end{pmatrix}$$

を $N$ 次**巡回行列**という．第 18.4 節で見るように巡回行列はディジタル信号処理におけるフィルタの理論と密接に関連している．

**定理 7.3**　$N$ 次巡回行列 $C = \mathrm{circ}\,(x_0, \cdots, x_{N-1})$ の行列式は, $\omega = \exp\left(\frac{2\pi i}{N}\right)$ として

$$|C| = \prod_{j=0}^{N-1} \left( \sum_{k=0}^{N-1} \omega^{jk} x_k \right)$$

である．

**証明.** $x_{k+lN} = x_k \ (k = 0, \cdots, N-1,\ l \in \boldsymbol{Z})$ とおく．

$$\omega^{k+lN} = \omega^k, \quad k = 0, \cdots, N-1,\ l \in \boldsymbol{Z}$$

であることに注意しておく．$k' = k - m$ とおくと

$$\begin{aligned} \sum_{k=0}^{N-1} \omega^{jk} x_{k-m} &= \sum_{k'=-m}^{N-1-m} \omega^{j(k'+m)} x_{k'} \\ &= \omega^{jm} \sum_{k'=-m}^{N-1-m} \omega^{jk'} x_{k'} = \omega^{jm} \sum_{k=0}^{N-1} \omega^{jk} x_k \end{aligned}$$

が成り立っている．したがって

$$U = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

とすると,

$$CU = \begin{pmatrix} \sum_{k=0}^{N-1} x_k & \sum_{k=0}^{N-1} x_k \omega^k & \cdots & \sum_{k=0}^{N-1} x_k \omega^{(N-1)k} \\ \sum_{k=0}^{N-1} x_{k-1} & \sum_{k=0}^{N-1} x_{k-1} \omega^k & \cdots & \sum_{k=0}^{N-1} x_{k-1} \omega^{(N-1)k} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} x_{k-N+1} & \sum_{k=0}^{N-1} x_{k-N+1} \omega^k & \cdots & \sum_{k=0}^{N-1} x_{k-N+1} \omega^{(N-1)k} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=0}^{N-1} x_k & \sum_{k=0}^{N-1} x_k \omega^k & \cdots & \sum_{k=0}^{N-1} x_k \omega^{(N-1)k} \\ \sum_{k=0}^{N-1} x_k & \omega \sum_{k=0}^{N-1} x_k \omega^k & \cdots & \omega^{N-1} \sum_{k=0}^{N-1} x_k \omega^{(N-1)k} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} x_k & \omega^{N-1} \sum_{k=0}^{N-1} x_k \omega^k & \cdots & \omega^{(N-1)(N-1)} \sum_{k=0}^{N-1} x_k \omega^{(N-1)k} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

$$\times \begin{pmatrix} \sum_{k=0}^{N-1} x_k & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{N-1} x_k \omega^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{k=0}^{N-1} x_k \omega^{(N-1)k} \end{pmatrix}$$

である．ただし，ここで最後の行の $\times$ は行列の積を表すものとする．ゆえに

$$|C|\,|U| = |U| \prod_{j=0}^{N-1} \left( \sum_{k=0}^{N-1} \omega^{jk} x_k \right)$$

である．$U$ はファンデルモンド行列であるから，定理 7.1 を用いて $|U| \neq 0$ が示せる．これより定理が成り立つことがわかる．■

## 7.4　固有多項式

$N$ 次正方行列 $A = (a_{ij})$ に対して，行列式

$$\Delta_A(z) = \det(zI_N - A) = \begin{vmatrix} z - a_{11} & -a_{12} & \cdots & -a_{1N} \\ -a_{21} & z - a_{22} & \cdots & -a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{N1} & -a_{N2} & \cdots & z - a_{NN} \end{vmatrix}$$

は $z$ の $N$ 次多項式になっている．これを $A$ の**固有多項式**あるいは**特性多項式**という．固有多項式は後に固有値に関する問題を調べるときに重要な役割をはたす．

$$\Delta_A(z) = a_0 z^N + a_1 z^{N-1} + \cdots + a_{N-1} z + a_N$$

とすると，

$$\Delta_A(z) = (z - a_{11}) \cdots (z - a_{NN}) + (z \text{ の } N-2 \text{ 次の多項式})$$

であるから，

$$\Delta_A(z) = z^N - (a_{11} + a_{22} + \cdots + a_{NN}) z^{N-1} + (z \text{ の } N-2 \text{ 次の多項式})$$

と表せている．したがって

$$\begin{aligned} a_0 &= 1 \\ a_1 &= -(a_{11} + a_{22} + \cdots + a_{NN}) \end{aligned}$$

が成り立っている．また $a_N = \Delta_A(0) = (-1)^N |A|$ であるから，

$$a_N = (-1)^N \det(A)$$

である．

$$\mathrm{tr}(A) = \sum_{k=1}^{N} a_{kk}$$

とし，これを $A$ の**トレース**という．

$$a_1 = -\mathrm{tr}(A)$$

である．

**定理 7.4**　$A$ を $N$ 次正方行列とし，$P$ を正則な $N$ 次正方行列とすると

$$\Delta_{P^{-1}AP}(x) = \Delta_A(z)$$

である．したがって，

$$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(A).$$

**証明.**

$$\begin{aligned}\Delta_{P^{-1}AP}(z) &= \det\left(zI_N - P^{-1}AP\right) = \det\left(P^{-1}(zI_N)P - P^{-1}AP\right)\\ &= \det\left(P^{-1}(zI_N - A)P\right) = \det(P)^{-1}\det(zI - A)\det(P)\\ &= \det(zI - A) = \Delta_A(z)\end{aligned}$$

である．これより $\Delta_{P^{-1}AP}(z)$ の係数と $\Delta_A(z)$ の係数は一致する．よって定理が証明された．■

## 7.5　小行列と小行列式

$(M, N)$ 行列 $A$ の成分の一部分を取り出して作った正方行列を $A$ の小行列といい，小行列の行列式を小行列式という．

小行列を扱うために記号を準備しておく．$L$ を自然数とする．$k$ を $k \leq L$ なる自然数とし，

$$\mathcal{I}_k(L) = \{(l_1, \cdots, l_k) : l_1, \cdots, l_k \in \boldsymbol{Z},\ 1 \leq l_1 < \cdots < l_k \leq L\}$$

とする．$M, N$ を自然数とし，$k, l$ を $1 \leq k \leq M$, $1 \leq l \leq N$ なる自然数とする．$J_1 = (m_1, \cdots, m_k) \in \mathcal{I}_k(M)$, $J_2 = (n_1, \cdots, n_l) \in \mathcal{I}_l(N)$ とする．$A = (a_{mn}) \in \boldsymbol{K}^{M\times N}$ に対して

$$A(J_1, J_2) = \begin{pmatrix} a_{m_1,n_1} & \cdots & a_{m_1,n_l} \\ \vdots & \ddots & \vdots \\ a_{m_k,n_1} & \cdots & a_{m_k,n_l} \end{pmatrix} \in \boldsymbol{K}^{k\times l}$$

と表す．$k = l$ の場合, 行列 $A(J_1, J_2)$ を $A$ の $k$ **次小行列**という．また $\det A(J_1, J_2)$ を $k$ **次小行列式**という．

特に $J_1 \in \mathcal{I}_k(M)$, $J_2 = \{n\} \in \mathcal{I}_1(N)$ の場合，$A(J_1, J_2)$ を $A(J_1, n)$ と表す．$A(m, J_2)$ も同様にして定義する．

**定理 7.5**　$A = (a_{ij}) \in \boldsymbol{K}^{M\times N}$, $B = (b_{nm}) \in \boldsymbol{K}^{N\times L}$ とし，$C = AB$ とする．

$k$ を $k \le \min\{M, L\}$ をみたす自然数とし，$J_1 = (m_1, \cdots, m_k) \in \mathcal{I}_k(M)$, $J_2 = (n_1, \cdots, n_k) \in \mathcal{I}_k(L)$ とする．このとき

$$C(J_1, J_2) = \sum_{m=1}^{N} A(J_1, m) B(m, J_2)$$

である．また

$$|C(J_1, J_2)| = \sum_{1 \le j_1 < \cdots < j_k \le N} \begin{vmatrix} a_{m_1, j_1} & \cdots & a_{m_1, j_k} \\ \vdots & \ddots & \vdots \\ a_{m_k, j_1} & \cdots & a_{m_k, j_k} \end{vmatrix} \begin{vmatrix} b_{j_1, n_1} & \cdots & b_{j_1, n_k} \\ \vdots & \ddots & \vdots \\ b_{j_k, n_1} & \cdots & b_{j_k, n_k} \end{vmatrix}$$

である．ただし，ここで $\sum_{1 \le j_1 < j_2 < \cdots < j_k \le N}$ は $1 \le j_1 < j_2 < \cdots < j_k \le N$ をみたすすべての $j_1, \cdots, j_k$ にわたってとる和を表す．

**証明.** コロン記法 (第 2.1 節参照) を用いる．行列の積の定義より

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} \begin{pmatrix} B(1,:) \\ \vdots \\ B(N,:) \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} B(1,:) + \cdots + a_{1N} B(N,:) \\ \vdots \\ a_{M1} B(1,:) + \cdots + a_{MN} B(N,:) \end{pmatrix}$$

$$= A(:,1) B(1,:) + \cdots + A(:,N) B(N,:).$$

したがって，

$$C(I_k, J_k) = A(I_k, 1) B(1, J_k) + \cdots + A(I_k, N) B(N, J_k)$$
$$= \sum_{m=1}^{N} A(I_k, m) B(m, J_k)$$

が成り立つ．行列式の性質 (I), (II) より

$$\det C(I_k, J_k) = \det \begin{pmatrix} a_{m_1,1} B(1, J_k) + \cdots + a_{m_1,N} B(N, J_k) \\ \vdots \\ a_{m_k,1} B(1, J_k) + \cdots + a_{m_k,N} B(N, J_k) \end{pmatrix}$$

$$= \sum_{j_1=1}^{N} a_{m_1,j_1} \det \begin{pmatrix} B(j_1, J_k) \\ a_{m_2,1}B(1,J_k) + \cdots + a_{m_2,N}B(N,J_k) \\ \vdots \\ a_{m_k,1}B(1,J_k) + \cdots + a_{m_k,N}B(N,J_k) \end{pmatrix}$$

$$= \cdots = \sum_{j_1=1}^{N} \cdots \sum_{j_k=1}^{N} a_{m_1,j_1} \cdots a_{m_k,j_k} \det \begin{pmatrix} B(j_1, J_k) \\ \vdots \\ B(j_k, J_k) \end{pmatrix} \quad (:= (i)\ \text{とおく}).$$

行列式の性質 (III), (IV) より

$$(i) = \sum_{1 \le j_1 < \cdots < j_k \le N} \sum \mathrm{sgn} \begin{pmatrix} j_1 & \cdots & j_k \\ j'_1 & \cdots & j'_k \end{pmatrix} a_{m_1,j'_1} \cdots a_{m_k,j'_k}$$

$$\times \begin{vmatrix} b_{j_1,n_1} & \cdots & b_{j_1,n_k} \\ \vdots & \ddots & \vdots \\ b_{j_k,n_1} & \cdots & b_{j_k,n_k} \end{vmatrix},$$

ただし，ここで $\sum$ は $\{j_1, \cdots, j_k\}$ の置換全体にわたってとる和を表すものとする．したがって

$$(i) = \sum_{1 \le j_1 < \cdots < j_k \le N} \begin{vmatrix} a_{m_1,j_1} & \cdots & a_{m_1,j_k} \\ \vdots & \ddots & \vdots \\ a_{m_k,j_1} & \cdots & a_{m_k,j_k} \end{vmatrix} \begin{vmatrix} a_{j_1,n_1} & \cdots & a_{j_1,n_k} \\ \vdots & \ddots & \vdots \\ a_{j_k,n_1} & \cdots & a_{j_k,n_k} \end{vmatrix}$$

を得る． ■

$(M, N)$ 行列 $A$ と $(N, M)$ 行列 $B$ の積 $AB$ は $M$ 次正方行列であるから，$AB$ の行列式が定義できる．その値は $A$ と $B$ の小行列式から次のようにして求まる．

**定理 7.6** $A = (a_{ij}) \in \boldsymbol{K}^{M \times N}$, $B = (b_{nm}) \in \boldsymbol{K}^{N \times M}$ とする．

(1) $M < N$ の場合,

$$|AB| = \sum_{J \in \mathcal{I}_M(N)} \det A(:, J) \det B(J, :)$$

$$= \sum_{1\le\mu_1<\mu_2<\cdots<\mu_M\le N} \begin{vmatrix} a_{1,\mu_1} & \cdots & a_{1,\mu_M} \\ \vdots & \ddots & \vdots \\ a_{M,\mu_1} & \cdots & a_{M,\mu_M} \end{vmatrix} \begin{vmatrix} b_{\mu_1,1} & \cdots & b_{\mu_1,M} \\ \vdots & \ddots & \vdots \\ b_{\mu_M,1} & \cdots & b_{\mu_M,M} \end{vmatrix},$$

ただしここで $\displaystyle\sum_{1\le\mu_1<\mu_2<\cdots<\mu_M\le N}$ は $1 \le \mu_1 < \mu_2 < \cdots < \mu_M \le N$ をみたすすべての $\mu_1, \cdots, \mu_M$ にわたってとる和を表す.

(2)　$M = N$ の場合，$|AB| = |A|\,|B|$ .

(3)　$M > N$ の場合，$|AB| = 0.$

**証明.** (1) は定理 7.5 の特別な場合である．(2) はすでに証明してある．(3) を示す.

$\mathbf{a}_j = A(:, j) \in \boldsymbol{K}^M$ として，$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ と表す.

$$AB = \begin{pmatrix} \sum_{n=1}^{N} a_{1n}b_{n1} & \cdots & \sum_{n=1}^{N} a_{1n}b_{nM} \\ \vdots & \ddots & \vdots \\ \sum_{n=1}^{N} a_{Mn}b_{n1} & \cdots & \sum_{n=1}^{N} a_{Mn}b_{nM} \end{pmatrix}$$

$$= \left( \sum_{n=1}^{N} b_{n1}\mathbf{a}_n, \sum_{n=1}^{N} b_{n2}\mathbf{a}_n, \cdots, \sum_{n=1}^{N} b_{nM}\mathbf{a}_n \right).$$

したがって,

$$|AB| = D\left( \sum_{n=1}^{N} b_{n1}\mathbf{a}_n, \sum_{n=1}^{N} b_{n2}\mathbf{a}_n, \cdots, \sum_{n=1}^{N} b_{nM}\mathbf{a}_n \right)$$

$$= \sum_{\mu_1,\cdots,\mu_M=1}^{N} D(\mathbf{a}_{\mu_1}, \cdots, \mathbf{a}_{\mu_M})\, b_{\mu_1,1} \cdots b_{\mu_M,M}.$$

ここで，$\mu_1, \cdots, \mu_M$ の中に同じ番号が含まれていれば，$D(\mathbf{a}_{\mu_1}, \cdots, \mathbf{a}_{\mu_M}) = 0$ であるが，$M > N$ なので必ず同じ番号は含まれていなければならない．よって (3) が証明された．■

この定理の応用として，グラム行列と呼ばれる行列の行列式を求める公式を示しておこう．グラム行列の応用例として本書では後に最良近似の問題を論ずる (第 10.5 節参照).

**定義 7.1**　$\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ に対して

$$G(\mathbf{a}_1, \cdots, \mathbf{a}_s) = \begin{pmatrix} \mathbf{a}_1^*\mathbf{a}_1 & \cdots & \mathbf{a}_1^*\mathbf{a}_s \\ \vdots & \ddots & \vdots \\ \mathbf{a}_s^*\mathbf{a}_1 & \cdots & \mathbf{a}_s^*\mathbf{a}_s \end{pmatrix} \in \boldsymbol{K}^{s\times s}$$

を $\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ のグラム行列という．

**系 7.7** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ とする．

(1) $A = (\mathbf{a}_1, \cdots, \mathbf{a}_s) \in \boldsymbol{K}^{N\times s}$ とすると

$$G(\mathbf{a}_1, \cdots, \mathbf{a}_s) = A^*A$$

である．

(2) $N < s$ のとき

$$\det G(\mathbf{a}_1, \cdots, \mathbf{a}_s) = 0$$

である．

(3) $1 \leq s \leq N$ のとき，$\mathbf{a}_1 = (a_{j1})_{j=1,\cdots,N}, \cdots, \mathbf{a}_s = (a_{js})_{j=1,\cdots,N}$ とすると

$$\det G(\mathbf{a}_1, \cdots, \mathbf{a}_s) = \sum_{1\leq\mu_1<\mu_2<\cdots<\mu_s\leq N} \left| \det \begin{pmatrix} a_{\mu_1,1} & \cdots & a_{\mu_1,s} \\ \vdots & \ddots & \vdots \\ a_{\mu_s,1} & \cdots & a_{\mu_s,s} \end{pmatrix} \right|^2$$

である．

**証明.** (1) 明らかに

$$A^*A = \begin{pmatrix} \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_s^* \end{pmatrix} (\mathbf{a}_1, \cdots, \mathbf{a}_s) = \begin{pmatrix} \mathbf{a}_1^*\mathbf{a}_1 & \cdots & \mathbf{a}_1^*\mathbf{a}_s \\ \vdots & \ddots & \vdots \\ \mathbf{a}_s^*\mathbf{a}_1 & \cdots & \mathbf{a}_s^*\mathbf{a}_s \end{pmatrix} \in \boldsymbol{K}^{s\times s}$$

である．定理 7.6 より $N < s$ ならば $\det(A^*A) = 0$ より (2) が成り立つ．また $1 \leq s \leq N$ の場合は定理 7.6 より

$$\det G(\mathbf{a}_1, \cdots, \mathbf{a}_s) = \det \left( \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{N1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1s}} & \cdots & \overline{a_{Ns}} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{Ns} \end{pmatrix} \right)$$

$$
= \sum_{1 \le \mu_1 < \mu_2 < \cdots < \mu_s \le N} \det \begin{pmatrix} \overline{a_{\mu_1,1}} & \cdots & \overline{a_{\mu_s,1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{\mu_1,s}} & \cdots & \overline{a_{\mu_s,s}} \end{pmatrix} \det \begin{pmatrix} a_{\mu_1,1} & \cdots & a_{\mu_1,s} \\ \vdots & \ddots & \vdots \\ a_{\mu_s,1} & \cdots & a_{\mu_s,s} \end{pmatrix}
$$

$$
= \sum_{1 \le \mu_1 < \mu_2 < \cdots < \mu_s \le N} \left| \det \begin{pmatrix} a_{\mu_1,1} & \cdots & a_{\mu_1,s} \\ \vdots & \ddots & \vdots \\ a_{\mu_s,1} & \cdots & a_{\mu_s,s} \end{pmatrix} \right|^2 .
$$

■

# 第8章　ブロック行列

行列をいくつかのブロックに分けて考えることを，行列のブロック化という．行列をブロック化して考えると，行列の計算の見通しがよくなることがある．また応用上は，いくつかのブロックからなる行列が自然に現れることもある．たとえば本書では第 19 章で述べるウェーブレットにその例が見出せる．本章ではブロック化による行列の計算方法について述べる．

**ポイント 8**　ブロック行列の計算では，特に定理 8.3 を使えるようにしておいてほしい．

## 8.1　ブロック行列の演算

ブロック化とは，たとえば

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 2 & 2 & 3 \\ 1 & 5 & 7 & 9 & 1 \\ 2 & 4 & 6 & 7 & 1 \end{pmatrix} \quad を \quad A = \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 2 & 2 & 3 \\ \hline 1 & 5 & 7 & 9 & 1 \\ 2 & 4 & 6 & 7 & 1 \end{array}\right)$$

のように区分けすることである．これは言いかえれば $A$ を

$$A_{11} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 2 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 9 & 1 \\ 7 & 1 \end{pmatrix}$$

という小さい行列からなる行列

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

とみなすことである．

一般に $(M, N)$ 行列 $A$ が $(m_i, n_j)$ 行列 $A_{ij}$, $m_1 + \cdots + m_p = M$, $n_1 + \cdots + n_q = N$ により

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix} \tag{8.1}$$

のようにブロック化されているとき，$A_{ij}$ を $A$ の $(i, j)$ **ブロック**という．また $A$ は $(m_i, n_j)$ 行列 $A_{ij}$ をブロックとする $(p, q)$ **ブロック行列**という．

$A$ を $(A_{ij})_{i=1,\cdots,p\,;\,j=1,\cdots,q}$ あるいは単に $(A_{ij})$ と表す．

同じ型のブロック行列同士の和は，各ブロックの和になっている．実際，次の定理が成り立つことは容易に確認できるであろう．

**定理 8.1**　$(M, N)$ 行列 $A$ が $(m_i, n_j)$ 行列 $A_{ij}$ をブロックとする $(p, q)$ ブロック行列で，$(M, N)$ 行列 $B$ が $(m_i, n_j)$ 行列 $B_{ij}$ をブロックとする $(p, q)$ ブロック行列であるとする．このとき，

$$A + B = \begin{pmatrix} A_{11} + B_{11} & \cdots & A_{1q} + B_{1q} \\ \vdots & \ddots & \vdots \\ A_{p1} + B_{p1} & \cdots & A_{pq} + B_{pq} \end{pmatrix}$$

が成り立つ．

この定理を帰納的に繰り返し用いることにより，次の一般的な定理が得られる．

**定理 8.2**　$(M, N)$ 行列 $A^{(k)}$ が $(m_i, n_j)$ 行列 $A_{ij}^{(k)}$ をブロックとする $(p, q)$ ブロック行列であるとする $(k = 1, \cdots, r)$．このとき，

$$\sum_{k=1}^{r} A^{(k)} = \begin{pmatrix} \displaystyle\sum_{k=1}^{r} A_{11}^{(k)} & \cdots & \displaystyle\sum_{k=1}^{r} A_{1q}^{(k)} \\ \vdots & \ddots & \vdots \\ \displaystyle\sum_{k=1}^{r} A_{p1}^{(k)} & \cdots & \displaystyle\sum_{k=1}^{r} A_{pq}^{(k)} \end{pmatrix}$$

が成り立つ．

ブロック行列の積は各ブロックを用いて次のように計算できる.

**定理 8.3** $(M, N)$ 行列 $A$ が $(m_i, n_j)$ 行列 $A_{ij}$ をブロックとする $(p, q)$ ブロック行列であり, また $B$ が $(N, L)$ 行列で, $(n_i, l_j)$ 行列 $B_{ij}$ をブロックとする $(q, r)$ ブロック行列, すなわち

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & \ddots & \vdots \\ B_{q1} & \cdots & B_{qr} \end{pmatrix}$$

とする. このとき, $AB$ は $(m_i, l_j)$ 行列 $\sum_{\mu=1}^{q} A_{i\mu}B_{\mu j}$ をブロックとする $(p, r)$ ブロック行列で,

$$AB = \begin{pmatrix} \sum_{\mu=1}^{q} A_{1\mu}B_{\mu 1} & \cdots & \sum_{\mu=1}^{q} A_{1\mu}B_{\mu r} \\ \vdots & \ddots & \vdots \\ \sum_{\mu=1}^{q} A_{p\mu}B_{\mu 1} & \cdots & \sum_{\mu=1}^{q} A_{p\mu}B_{\mu r} \end{pmatrix}$$

が成り立っている.

**証明.** $A = (a_{jk})_{\substack{j=1,\cdots,M \\ k=1,\cdots,N}}$, $B = (b_{jk})_{\substack{j=1,\cdots,N \\ k=1,\cdots,L}}$ とおく. まず $q = 1$ の場合を証明する. このとき $A$ は $(M, n_1)$ 行列, $B$ は $(n_1, L)$ 行列である.

$$\begin{pmatrix} A_{11} \\ \vdots \\ A_{p1} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1r} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & \ddots & \vdots \\ C_{p1} & \cdots & C_{pr} \end{pmatrix}$$

とおく. ただし $C_{ij}$ は $(m_i, l_j)$ 行列である. このとき, $C_{ij}$ の $(s, t)$ 成分 ( $s = 1, \cdots, m_i$, $t = 1, \cdots, l_j$ ) は $AB$ の第 $(M_i + s, L_j + t)$ 成分である. ただし, ここで $M_1 = 0$, $L_1 = 0$, また $i > 1$, $j > 1$ に対しては

$$M_i = m_1 + \cdots + m_{i-1}, \quad L_j = l_1 + \cdots + l_{j-1}$$

とする. したがって, $C_{ij}$ の $(s, t)$ 成分は

$$\sum_{k=1}^{n_1} a_{M_i+s,k}\, b_{k,L_j+t} \tag{8.2}$$

である. 一方, $a_{M_i+s,k}$ は $A_{i1}$ の $(s, k)$ 成分であり, $b_{k,L_j+t}$ は $B_{1j}$ の $(k, t)$ 成分であるから, (8.2) は $A_{i1}B_{1j}$ の $(s, t)$ 成分を表している. したがって $C_{ij} = A_{i1}B_{1j}$

が得られる．ゆえに

$$\begin{pmatrix} A_{11} \\ \vdots \\ A_{p1} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1r} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & \cdots & A_{11}B_{1r} \\ \vdots & \ddots & \vdots \\ A_{p1}B_{11} & \cdots & A_{p1}B_{1r} \end{pmatrix}$$

である．次に一般の場合を示す．

$$A_j = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{pj} \end{pmatrix}, \quad B_j = \begin{pmatrix} B_{j1} & \cdots & B_{jr} \end{pmatrix}, \quad j = 1, \cdots, q$$

とおくと

$$A = \begin{pmatrix} A_1 & \cdots & A_q \end{pmatrix}, \quad B = \begin{pmatrix} B_1 \\ \vdots \\ B_q \end{pmatrix}$$

と表せる．$A_j$ は $(M, n_j)$ 行列，$B_j$ は $(n_j, L)$ 行列である．$AB$ の $(m,l)$ 成分 ( $m = 1, \cdots, M$, $l = 1, \cdots, L$ ) は

$$\sum_{\mu=1}^{N} a_{m\mu}b_{\mu l} = \sum_{\mu=1}^{n_1} a_{m\mu}b_{\mu l} + \sum_{\mu=n_1+1}^{n_1+n_2} a_{m\mu}b_{\mu l} + \cdots + \sum_{\mu=n_1+\cdots+n_{q-1}+1}^{n_1+\cdots+n_q} a_{m\mu}b_{\mu l}$$

であり，ここで $\sum_{\mu=1}^{n_1} a_{m\mu}b_{\mu l}$ は $A_1B_1$ の $(m,l)$ 成分，$\sum_{\mu=n_1+1}^{n_1+n_2} a_{m\mu}b_{\mu l}$ は $A_2B_2$ の $(m,l)$ 成分，$\cdots$，そして $\sum_{\mu=n_1+\cdots+n_{q-1}+1}^{n_1+\cdots+n_q} a_{m\mu}b_{\mu l}$ は $A_qB_q$ の $(m,l)$ 成分である．したがって $\sum_{\mu=1}^{N} a_{m\mu}b_{\mu l}$ は $\sum_{j=1}^{q} A_jB_j$ の $(m,l)$ 成分である．ゆえに

$$AB = \sum_{j=1}^{q} A_jB_j$$

となっている．ここで各 $A_jB_j$ はすでに証明した $q=1$ の場合の議論から，

$$A_jB_j = \begin{pmatrix} A_{1j}B_{j1} & \cdots & A_{1j}B_{jr} \\ \vdots & \ddots & \vdots \\ A_{pj}B_{j1} & \cdots & A_{pj}B_{jr} \end{pmatrix}.$$

よって，前定理より

$$AB = \sum_{j=1}^{q} A_j B_j = \begin{pmatrix} \sum_{j=1}^{q} A_{1j}B_{j1} & \cdots & \sum_{j=1}^{q} A_{1j}B_{jr} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{q} A_{pj}B_{j1} & \cdots & \sum_{j=1}^{q} A_{pj}B_{jr} \end{pmatrix}$$

を得る． ■

**例 8.1** $A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$, $B = \begin{pmatrix} B_{11} & B_{12} \\ O' & B_{22} \end{pmatrix}$ とする．ここで $A_{ij}$ は $(m_i, n_j)$ 行列で，$O$ は $(m_2, n_1)$ 零行列，また $B_{ij}$ は $(n_i, l_j)$ 行列，$O'$ は $(n_2, l_1)$ 零行列とする．このとき

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}O' & A_{11}B_{12} + A_{12}B_{22} \\ OB_{11} + A_{22}O' & OB_{12} + A_{22}B_{22} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} + A_{12}B_{22} \\ O_{m_2 \times l_1} & A_{22}B_{22} \end{pmatrix}$$

である．

ブロック行列の応用として次のことが証明できる．

**定理 8.4** $A = (a_{jk})$ を $N$ 次正方行列で，正則であるとする．$A$ が上三角行列ならば $A^{-1}$ も上三角行列である．また $A$ が下三角行列ならば $A^{-1}$ も下三角行列である．いずれの場合も $a_{kk} \neq 0$, $k = 1, \cdots, N$ であり，$A^{-1}$ の $(k, k)$ 成分は $\dfrac{1}{a_{kk}}$ である．

**証明.** 1 次正方行列の場合は明らかである．$n$ 次正方行列の場合に正しいとして，$n+1$ 次の場合を示す．

$$A = \begin{pmatrix} P & Q \\ O_{1\times n} & s \end{pmatrix}, \qquad \begin{matrix} P \in \boldsymbol{K}^{n\times n},\ s \in \boldsymbol{K}, \\ Q \in \boldsymbol{K}^{n\times 1} \end{matrix}$$

とブロック化する．同様に $A^{-1} = \begin{pmatrix} X & Y \\ W & z \end{pmatrix}$ も $A$ と同じタイプのブロック化をする．このとき，

$$\begin{pmatrix} I_n & O_{n\times 1} \\ O_{1\times n} & 1 \end{pmatrix} = \begin{pmatrix} P & Q \\ O_{1\times n} & s \end{pmatrix} \begin{pmatrix} X & Y \\ W & z \end{pmatrix}$$
$$= \begin{pmatrix} PX + QW & PY + Qz \\ sW & sz \end{pmatrix}$$

である．$sz = 1$ より $s \neq 0$ である．ゆえに $sW = O_{1\times n}$ より $W = O_{1\times n}$ が得られる．したがって，

$$I_n = PX + QW = PX$$

となり，$X = P^{-1}$ である (系 6.4 参照)．帰納法の仮定から $X$ は上三角行列である．ゆえに $A^{-1} = \begin{pmatrix} X & Y \\ O_{1\times n} & z \end{pmatrix}$ も上三角行列である．下三角行列の場合も同様に示せる．$A, A^{-1}$ はともに上 (下) 三角行列であるから，$A^{-1} = (b_{kj})$ とすると $AA^{-1}$ の対角成分は $a_{kk}b_{kk}\ (k = 1, \cdots, n)$ であり，これらは 1 である．■

## 8.2　ブロック行列の逆行列

行列 $A_0$ の逆行列がわかっているときに，$A_0$ に行と列を加えて拡大した行列 $A$ の逆行列を，$A_0$ に関する計算結果を援用して求める方法がある．それを次に述べよう．

**定理 8.5**　$A \in \boldsymbol{K}^{N\times N}$ を

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

ただし, $A_{11} \in \boldsymbol{K}^{r\times r}$, $A_{22} \in \boldsymbol{K}^{(N-r)\times(N-r)}$, $A_{12} \in \boldsymbol{K}^{r\times(N-r)}$, $A_{21} \in \boldsymbol{K}^{(N-r)\times r}$ とする．$A_{11}$ が正則であるとする．このとき, $A$ が正則であることと, $-A_{21}A_{11}^{-1}A_{12}+A_{22}$ が正則であることは同値である．このとき

$$B = -A_{21}A_{11}^{-1}A_{12} + A_{22}$$

とおくと，$A$ の逆行列は

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}B^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B^{-1} \\ -B^{-1}A_{21}A_{11}^{-1} & B^{-1} \end{pmatrix}$$

により与えられる．

**証明.** $A$ が正則であるとする．このとき，$A^{-1}$ を $A$ と同じ形にブロック化し，それを

$$A^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

とおく．$AA^{-1} = I_N$ であるから，

$$\begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & I_{N-r} \end{pmatrix}$$

となる．仮定より $A_{11}$ が正則であり，$A_{11}B_{12} + A_{12}B_{22} = O$ であるから

$$B_{12} = -A_{11}^{-1}A_{12}B_{22} \tag{8.3}$$

である．したがって，

$$\begin{aligned} I_{N-r} &= A_{21}B_{12} + A_{22}B_{22} = -A_{21}A_{11}^{-1}A_{12}B_{22} + A_{22}B_{22} \\ &= \left(-A_{21}A_{11}^{-1}A_{12} + A_{22}\right)B_{22}. \end{aligned}$$

このことから，$B = -A_{21}A_{11}^{-1}A_{12} + A_{22}$ は正則であり，$B_{22} = B^{-1}$ となっている．ゆえに (8.3) より

$$B_{12} = -A_{11}^{-1}A_{12}B^{-1}$$

が得られる．ところで $A^{-1}A = I_N$ でもあるから

$$\begin{pmatrix} B_{11}A_{11} + B_{12}A_{21} & B_{11}A_{12} + B_{12}A_{22} \\ B_{21}A_{11} + B_{22}A_{21} & B_{21}A_{12} + B_{22}A_{22} \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & I_{N-r} \end{pmatrix}$$

である．したがって，

$$B_{21} = -B_{22}A_{21}A_{11}^{-1} = -B^{-1}A_{21}A_{11}^{-1}.$$

また $B_{11}A_{11} + B_{12}A_{21} = I_r$ より

$$\begin{aligned} B_{11} &= \left(I_r - B_{12}A_{21}\right)A_{11}^{-1} \\ &= A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1} \end{aligned}$$

が得られる．以上より，

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}B^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B^{-1} \\ -B^{-1}A_{21}A_{11}^{-1} & B^{-1} \end{pmatrix} \tag{8.4}$$

が導かれる．

逆に，$B = -A_{21}A_{11}^{-1}A_{12} + A_{22}$ が正則であるとき，(8.4) で定義される行列が $A$ の逆行列になっていることは直接，計算をして容易に証明される．■

**例題 8.1** $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ の逆行列は $\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$ である．この結果を用いて，$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ の逆行列を求めよ．

**答.** $A_{11} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, $A_{12} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $A_{21} = (1 \quad 0)$, $A_{22} = 1$ とブロック分けする．$A_{12}$ は零行列であるから，$B = -A_{21}A_{11}^{-1}A_{12} + A_{22} = A_{22} = 1$, $B_{12} = O$, ゆえに $B_{22} = 1$，さらに

$$B_{21} = -\begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \end{pmatrix},$$

$$B_{11} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}.$$

したがって，$A^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$. ■

定理 8.5 と同様に次も成り立つ．

**定理 8.6** $A \in \boldsymbol{K}^{N \times N}$ を

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

ただし, $A_{11} \in \boldsymbol{K}^{r \times r}$, $A_{22} \in \boldsymbol{K}^{(N-r) \times (N-r)}$, $A_{12} \in \boldsymbol{K}^{r \times (N-r)}$, $A_{21} \in \boldsymbol{K}^{(N-r) \times r}$ とする．$A_{22}$ が正則であるとする．このとき，$A$ が正則であることと，$A_{11} - A_{12}A_{22}^{-1}A_{21}$ が正則であることは同値であり，$B = A_{11} - A_{12}A_{22}^{-1}A_{21}$ とすると

$$A^{-1} = \begin{pmatrix} B^{-1} & -B^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}B^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}B^{-1}A_{12}A_{22}^{-1} \end{pmatrix}$$

で与えられる．

**問題 8.1** $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ の逆行列は $\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$ である．この結果を用いて，$A =$

$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ の逆行列を求めよ．

## 8.3 ブロック行列の行列式

$A \in \boldsymbol{K}^{N\times N}$ とし，$A$ が次のようにブロック化されているとしよう．

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{8.5}$$

ただし，$A_{11} \in \boldsymbol{K}^{r\times r}$, $A_{12} \in \boldsymbol{K}^{r\times(N-r)}$, $A_{21} \in \boldsymbol{K}^{(N-r)\times r}$, $A_{22} \in \boldsymbol{K}^{(N-r)\times(N-r)}$ であるとする．このとき次のことが成り立つ．

**定理 8.7**　$A$ を (8.5) で定義されるブロック行列とする．

(1)　$A_{12} = O_{r\times(N-r)}$ または $A_{21} = O_{(N-r)\times r}$ ならば

$$|A| = |A_{11}|\,|A_{22}|.$$

(2)　$A_{22}$ が正則ならば

$$|A| = |A_{22}|\,\left|A_{11} - A_{12}A_{22}^{-1}A_{21}\right|.$$

(3)　$A_{11}$ が正則ならば

$$|A| = |A_{11}|\,\left|A_{22} - A_{21}A_{11}^{-1}A_{12}\right|.$$

**証明.** (1)　$A_{12} = O_{r\times(N-r)}$ の場合を証明する．この場合

$$A = \begin{pmatrix} A_{11} & O_{r\times(N-r)} \\ A_{21} & I_{N-r} \end{pmatrix} \begin{pmatrix} I_{N_1} & O_{r\times(N-r)} \\ O_{(N-r)\times r} & A_{22} \end{pmatrix}$$

が成り立っている．$k = 1, \cdots, N-r-1$ に対して

$$A_{21}^{(k)} = A_{21}\,(1 : N-r-k, 1 : r)$$

と定義する．次式の左辺の行列式の第 $N$ 列に関する余因子展開をすれば

$$\begin{vmatrix} A_{11} & O_{r\times(N-r)} \\ A_{21} & I_{N-r} \end{vmatrix} = \begin{vmatrix} A_{11} & O_{r\times(N-r-1)} \\ A_{21}^{(1)} & I_{N-r-1} \end{vmatrix} \tag{8.6}$$

が得られる．次に (8.6) の右辺を第 $N-1$ 列に関して余因子展開し，さらに同様の操作を続けることにより，

$$\begin{vmatrix} A_{11} & O_{r\times(N-r-1)} \\ A_{21}^{(1)} & I_{N-r-1} \end{vmatrix} = \begin{vmatrix} A_{11} & O_{r\times(N-r-2)} \\ A_{21}^{(2)} & I_{N-r-2} \end{vmatrix} = \cdots = |A_{11}|$$

が得られる．同様にして ( ただし今度は第 1 列から余因子展開していくようにすれば )

$$\left|\begin{pmatrix} I_{N_1} & O_{r\times(N-r)} \\ O_{(N-r)\times r} & A_{22} \end{pmatrix}\right| = |A_{22}|$$

も導かれる．したがって (1) が成り立つ．

(2)

$$\begin{pmatrix} I_r & -A_{12}A_{22}^{-1} \\ O_{(N-r)\times r} & I_{N-r} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_{N_1} & O_{r\times(N-r)} \\ -A_{22}^{-1}A_{21} & I_{N-r} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & O_{r\times(N-r)} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_{N_1} & O_{r\times(N-r)} \\ -A_{22}^{-1}A_{21} & I_{N-r} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & O_{r\times(N-r)} \\ O_{(N-r)\times r} & A_{22} \end{pmatrix}$$

である．

$$\begin{vmatrix} I_r & -A_{12}A_{22}^{-1} \\ O_{(N-r)\times r} & I_{N-r} \end{vmatrix} = 1, \quad \begin{vmatrix} I_r & O_{r\times(N-r)} \\ -A_{22}^{-1}A_{21} & I_{N-r} \end{vmatrix} = 1,$$

$$\begin{vmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & O_{r\times(N-r)} \\ O_{(N-r)\times r} & A_{22} \end{vmatrix} = |A_{22}|\,|A_{11} - A_{12}A_{22}^{-1}A_{21}|$$

(最後の等式は (1) による) であるから，

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |A_{22}|\,|A_{11} - A_{12}A_{22}^{-1}A_{21}|$$

が得られる．

(3)　次の等式から (2) と同様にして証明できる．

$$\begin{pmatrix} I_r & O_{r\times(N-r)} \\ -A_{21}A_{11}^{-1} & I_{N-r} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_r & -A_{11}^{-1}A_{12} \\ O_{(N-r)\times r} & I_{N-r} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} & O_{r\times(N-r)} \\ O_{(N-r)\times r} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

■

# 第2部

# 理論編

——線形構造と基底——

# 第9章　基底と部分空間

この章では，線形独立，線形従属，基底，$\boldsymbol{K}^N$ の線形部分空間，内積など，線形代数の理論的な根幹をなす概念を学ぶ．本章と次章は多少抽象的な話であるが，すぐ後でその抽象論が連立1次方程式の解法，多変量解析，離散フーリエ解析，離散ウェーブレット解析など具体的な応用の中に溶け込んでいくことが解説される．

**ポイント9**　この章で述べる概念はどれも線形代数で中心的な役割を果たすものである．線形独立，線形従属，線形包，部分空間，基底の概念は必ず把握しておくこと．

## 9.1　線形独立，線形従属

ベクトル $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ が与えられたとき，これらのベクトルの間に相互関係があるかどうかを知ることがしばしば必要になる．相互関係をどのように数学的に定式化するかは，考えている問題にもよるが，線形代数ではしばしば次のようなことが問題とされる．

> $\mathbf{a}_1, \cdots, \mathbf{a}_s$ の中に自分自身以外のベクトルの線形結合 (第1.2節参照) で表されるものがあるか？　すなわち，
>
> $$\mathbf{a}_j = c_1\mathbf{a}_1 + \cdots + c_{j-1}\mathbf{a}_{j-1} + c_{j+1}\mathbf{a}_{j+1} + \cdots + c_s\mathbf{a}_s \tag{9.1}$$
>
> の関係式をみたすある $\mathbf{a}_j$ が存在するか？

もし $\mathbf{a}_1, \cdots, \mathbf{a}_s$ の中で，(9.1) をみたすような $\mathbf{a}_j$ が存在するとき，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ は**線形従属の関係にある**，あるいは単に**線形従属**であるという．これは次のように言い換えることができる．

**定理 9.1** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ が線形従属であるとは,

$$\lambda_1 \mathbf{a}_1 + \cdots + \lambda_s \mathbf{a}_s = \mathbf{0} \tag{9.2}$$

をみたす $\lambda_1, \cdots, \lambda_s \in \boldsymbol{K}$ で $\lambda_1 = \cdots = \lambda_s = 0$ でないものが存在することである.

**証明.** $\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形従属ならば (9.1) の関係式が成り立つので,

$$c_1 \mathbf{a}_1 + \cdots + c_{j-1} \mathbf{a}_{j-1} - 1\mathbf{a}_j + c_{j+1} \mathbf{a}_{j+1} + \cdots + c_s \mathbf{a}_s = \mathbf{0}$$

が成り立つ. 逆に (9.2) をみたす $\lambda_1, \cdots, \lambda_s \in \boldsymbol{K}$ で $\lambda_1 = \cdots = \lambda_s = 0$ でないものが存在するならば, $\lambda_j \neq 0$ なる $\lambda_j$ をとることができるので,

$$\mathbf{a}_j = -\frac{\lambda_1}{\lambda_j} \mathbf{a}_1 - \cdots - \frac{\lambda_{j-1}}{\lambda_j} \mathbf{a}_{j-1} - \frac{\lambda_{j+1}}{\lambda_j} \mathbf{a}_{j+1} - \cdots - \frac{\lambda_s}{\lambda_j} \mathbf{a}_s$$

が得られる. ■

線形従属でないことを線形独立であるという. すなわち $\mathbf{a}_1, \cdots, \mathbf{a}_s$ が**線形独立**であるとは

$$\lambda_1 \mathbf{a}_1 + \cdots + \lambda_s \mathbf{a}_s = \mathbf{0}$$

をみたす $\lambda_1, \cdots, \lambda_s \in \boldsymbol{K}$ が $\lambda_1 = \cdots = \lambda_s = 0$ に限ることである.

線形従属のことを 1 **次従属**, また線形独立のことを 1 **次独立**ともいう.

たとえば, 明らかに $N$ 次元基本ベクトル $\mathbf{e}_1, \cdots, \mathbf{e}_N$ は線形独立である. なぜならば,

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} = \lambda_1 \mathbf{e}_1 + \cdots + \lambda_N \mathbf{e}_N$$

であるから, $\lambda_1 \mathbf{e}_1 + \cdots + \lambda_N \mathbf{e}_N = \mathbf{0}$ ならば $\lambda_1 = \cdots = \lambda_N = 0$ となるからである.

**命題 9.2** (1) $\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形独立ならば, その中の任意の $r$ 個 $(1 \leq r \leq s)$ のベクトルも線形独立になっている.

(2) $\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形従属ならば, それに $k$ 個 $(1 \leq k)$ の任意のベクトルを加えた $\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{a}_{s+1}, \cdots, \mathbf{a}_{s+k}$ も線形従属である.

(3) $A$ を $N$ 次正則行列とする. $\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形独立ならば, $A\mathbf{a}_1, \cdots, A\mathbf{a}_s$ も線形独立である.

**証明.** (1)　$1 \leq i_1 < \cdots < i_r \leq s$ とし，残りの番号を $\nu_1 < \cdots < \nu_{s-r}$ とする．$\lambda_{i_1}\mathbf{a}_{i_1} + \cdots + \lambda_{i_r}\mathbf{a}_{i_r} = \mathbf{0}$ とする．$\lambda_{\nu_1} = \cdots = \lambda_{\nu_{s-r}} = 0$ とすると $\lambda_1\mathbf{a}_1 + \cdots + \lambda_N\mathbf{a}_N = 0$ であるから，$\lambda_1 = \cdots = \lambda_N = 0$ となる．

(2)　ある $(\lambda_1, \cdots, \lambda_s) \neq (0, \cdots, 0)$ が存在し，$\lambda_1\mathbf{a}_1 + \cdots + \lambda_s\mathbf{a}_s = 0$ である．$\lambda_{s+1} = \cdots = \lambda_{s+k} = 0$ とすれば，$\lambda_1\mathbf{a}_1 + \cdots + \lambda_{s+k}\mathbf{a}_{s+k} = \mathbf{0}$ である．

(3)　$\lambda_1 A\mathbf{a}_1 + \cdots + \lambda_s A\mathbf{a}_s = \mathbf{0}$ とする．このとき $A\left(\sum_{j=1}^{s}\lambda_j\mathbf{a}_j\right) = \mathbf{0}$ であるが，$A^{-1}$ が存在するので，

$$\sum_{j=1}^{s}\lambda_j\mathbf{a}_j = A^{-1}A\left(\sum_{j=1}^{s}\lambda_j\mathbf{a}_j\right) = A^{-1}\mathbf{0} = \mathbf{0}$$

である．したがって $\lambda_1 = \cdots = \lambda_s = 0$ となる．■

**問題 9.1**　$\mathbf{a}_1, \cdots, \mathbf{a}_s$ の任意の二つのベクトル $\mathbf{a}_i$, $\mathbf{a}_j$ $(i \neq j)$ が線形独立であれば，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ も線形独立か？

$\boldsymbol{K}^N$ の $N$ 個のベクトルについて，次のことが成り立つ．

**定理 9.3**　$\mathbf{a}_1, \cdots, \mathbf{a}_N \in \boldsymbol{K}^N$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形独立であるための必要十分条件は，第 5.4.3 節で導入した記号を用いて表せば

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_N) \neq 0$$

である．したがって，$\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形従属であるための必要十分条件は

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_N) = 0$$

である．

**証明.** $\lambda_1\mathbf{a}_1 + \cdots + \lambda_N\mathbf{a}_N = \mathbf{0}$ は $\mathbf{a}_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{Nk} \end{pmatrix}$ $(k = 1, \cdots, N)$ とすると

$$\begin{cases} a_{11}\lambda_1 + \cdots + a_{1N}\lambda_N = 0 \\ \qquad\qquad\vdots \\ a_{N1}\lambda_1 + \cdots + a_{NN}\lambda_N = 0 \end{cases}$$

と表せる．したがって，$\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形独立であるとは，この斉次連立 1 次方程式が非自明な解をもたないことである．したがって定理 6.5 より，定理の前半の主張が示せた．後半の主張は前半の主張より明らかである．■

**系 9.4** (1) $\mathbf{a}_1, \cdots, \mathbf{a}_N \in \boldsymbol{K}^N$ が線形独立であるとする．このとき，任意の $\mathbf{b} \in \boldsymbol{K}^N$ は $\mathbf{a}_1, \cdots, \mathbf{a}_N$ の線形結合で表される．

(2) $N'$ 個のベクトル $\mathbf{a}_1, \cdots, \mathbf{a}_{N'} \in \boldsymbol{K}^N$ は $N' > N$ のときは，必ず線形従属になる．

**証明．** (1) $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．定理 9.3 より $|A| \neq 0$ であるから，$A^{-1}$ が存在する．任意の $\mathbf{b} \in \boldsymbol{K}^N$ に対して，

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} = A^{-1}\mathbf{b}$$

とおくと

$$\begin{aligned} \mathbf{b} = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} &= (\mathbf{a}_1, \cdots, \mathbf{a}_N) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix} \\ &= \lambda_1 \mathbf{a}_1 + \cdots + \lambda_N \mathbf{a}_N \end{aligned}$$

が得られる．

(2) $\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形従属の場合は明らか．もし $\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形独立の場合，(1) より

$$\mathbf{a}_{N+1} = \lambda_1 \mathbf{a}_1 + \cdots + \lambda_N \mathbf{a}_N$$

と表せる．したがって，$\mathbf{a}_1, \cdots, \mathbf{a}_N, \mathbf{a}_{N+1}$ は線形従属であるから，命題 9.2 (2) より (2) が示される．■

$N' < N$ の場合に，$\mathbf{a}_1, \cdots, \mathbf{a}_{N'} \in \boldsymbol{K}^N$ が線形独立か線形従属かを判別する方法もある．それは次節で，若干の準備の後に示す．

## 9.2 線形包

ベクトル $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ に対して，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ の線形結合で表されるベクトル全体のなす集合を $\operatorname{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ と表し，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ の**線形包**(linear span) という．すなわち

$$\operatorname{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\} = \left\{ \sum_{j=1}^{s} \lambda_j \mathbf{a}_j : \lambda_1, \cdots, \lambda_s \in \boldsymbol{K} \right\}$$

である．

系 9.4 (1) より，$\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形独立ならば

$$\boldsymbol{K}^N = \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$$

であることがわかる．じつはこの逆も次の定理が示すように成立する．

**定理 9.5**　$\mathbf{a}_1, \cdots, \mathbf{a}_N \in \boldsymbol{K}^N$ とする．次の (1) ∼ (4) は同値である．

(1)　$\boldsymbol{K}^N = \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$.

(2)　$\mathbf{a}_1, \cdots, \mathbf{a}_N$ が線形独立.

(3)　$D(\mathbf{a}_1, \cdots, \mathbf{a}_N) \neq 0$.

(4)　$(N, N)$ 行列 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ は正則行列である.

**証明.** (2) ⇒ (1) はすでに示した．まず (1) ⇒ (4) を証明する．仮定より各 $\mathbf{e}_j$ は $\mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ に属するので $\mathbf{e}_j = \lambda_{1j}\mathbf{a}_1 + \cdots + \lambda_{Nj}\mathbf{a}_N$ と表せる．したがって，

$$(\mathbf{a}_1, \cdots, \mathbf{a}_N)\begin{pmatrix} \lambda_{j1} \\ \vdots \\ \lambda_{jN} \end{pmatrix} = \mathbf{e}_j$$

であるから，

$$(\mathbf{a}_1, \cdots, \mathbf{a}_N)\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1N} \\ \vdots & \ddots & \vdots \\ \lambda_{N1} & \cdots & \lambda_{NN} \end{pmatrix} = (\mathbf{e}_1, \cdots, \mathbf{e}_N) = I_N$$

である．したがって，$(\mathbf{a}_1, \cdots, \mathbf{a}_N)$ は逆行列をもつ．ゆえに (4) が成り立つ．次に (3) ⇔ (2) は定理 9.3 で示している．また (3) ⇔ (4) は定理 6.3 に他ならない．

■

この定理において，ベクトル $\mathbf{a}_1, \cdots, \mathbf{a}_N$ が $N$ 個であるという仮定は本質的であることに注意する．

**定理 9.6**　$\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{b} \in \boldsymbol{K}^N$ とし，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形独立であるとする．このとき $\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{b}$ が線形従属であるための必要十分条件は

$$\mathbf{b} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$$

となることである．

**証明.** 十分性は定義より明らかである．必要性を証明する．線形従属という仮定より，ある $(\lambda_1, \cdots, \lambda_{s+1}) \neq (0, \cdots, 0)$ で $\lambda_1 \mathbf{a}_1 + \cdots + \lambda_s \mathbf{a}_s + \lambda_{s+1} \mathbf{b} = 0$ となるものが存在する．このとき，$\lambda_{s+1} \neq 0$ である．なぜなら，もし $\lambda_{s+1} = 0$ であるとすると，$\lambda_1 \mathbf{a}_1 + \cdots + \lambda_s \mathbf{a}_s = 0$ より $\lambda_1 = \cdots = \lambda_s$ となり，$(\lambda_1, \cdots, \lambda_{s+1}) \neq (0, \cdots, 0)$ に矛盾するからである．したがって

$$\mathbf{b} = -\sum_{j=1}^{s} \frac{\lambda_j}{\lambda_{s+1}} \mathbf{a}_j \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}.$$

■

**定理 9.7** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ とする．もしこの中の一つのベクトル $\mathbf{a}_k$ が他の残りの $s-1$ 個のベクトルの線形結合で表せているならば，

$$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\} = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \cdots, \mathbf{a}_s\}$$

である．

**証明.** $\mathbf{a}_k = \lambda_1 \mathbf{a}_1 + \cdots + \lambda_{k-1} \mathbf{a}_{k-1} + \lambda_{k+1} \mathbf{a}_{k+1} + \cdots + \lambda_s \mathbf{a}_s$ であるとする．任意の $\mathbf{x} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ に対して，

$$\begin{aligned} \mathbf{x} &= \alpha_1 \mathbf{a}_1 + \cdots + \alpha_s \mathbf{a}_s \\ &= \sum_{j=1}^{k-1} (\alpha_j + \alpha_k \lambda_j) \mathbf{a}_j + \sum_{j=k+1}^{s} (\alpha_j + \alpha_k \lambda_j) \mathbf{a}_j \\ &\in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \cdots, \mathbf{a}_s\} \end{aligned}$$

であるから，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\} \subset \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \cdots, \mathbf{a}_s\}$ である．"$\supset$"は明らかなので，定理が得られる．■

次の定理は理論上有用なものである．

**定理 9.8** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$, $s < N$ が線形独立であるとする．このとき，基本ベクトル $\mathbf{e}_1, \cdots, \mathbf{e}_N$ の中から $N-s$ 個 $\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-s}}$ を適当に選び，$\mathbf{a}_1, \cdots, \mathbf{a}_s$, $\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-s}}$ が線形独立で，

$$\boldsymbol{K}^N = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-s}}\}$$

となるようにできる．

**証明.** まず前半の主張を証明する．$\mathbf{a}_{s+1} = \mathbf{e}_1$, $\cdots$, $\mathbf{a}_{s+N} = \mathbf{e}_N$ とおく．このとき明らかに

$$\boldsymbol{K}^N = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{s+N}\}$$

である．$s+N>N$ であるから系 9.4 (2) より $\mathbf{a}_1, \cdots, \mathbf{a}_{s+N}$ は線形従属である．そこで $\mathbf{a}_1, \cdots, \mathbf{a}_t$ が線形従属になる最初の番号 $t$ を $t_1$ とおく．仮定より $s+1 \leq t_1 \leq s+N$ である．このとき，$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}$ は線形独立であるから，定理 9.6 より $\mathbf{a}_{t_1} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}\}$ である．したがって，定理 9.7 より

$$\boldsymbol{K}^N = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_{s+N}\}$$

が成り立つ．$s+N-1=N$ ならば，$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_{s+N}$ は $N$ 個のベクトルなので，定理 9.5 より線形独立であることがわかり，証明が終わる．一方，もし $s+N-1>N$ であれば，$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_{s+N}$ は線形従属である．そこで $\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_t$ が最初に線形従属になる番号 $t$ を $t_2$ とおく．$t_1+1 \leq t_2 \leq s+N$ である．$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_{s+N}$ から $\mathbf{a}_{t_2}$ を除いたベクトルの線形包は，上の同じ理由で $\boldsymbol{K}^N$ と一致する．ここでもし $s+N-2=N$ ならば，上と同様の理由で証明は終了する．$s+N-2>N$ の場合は上の議論を続ける．この議論を $s+N-k=N$ となるまで繰り返し続ければ，証明は終了する．■

この定理を用いて $s$ 個のベクトル $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ の線形独立性を判定する方法を示す．

**定理 9.9**　$1 \leq s \leq N$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ とし，$A = (\mathbf{a}_1, \cdots, \mathbf{a}_s) = (a_{ij})_{\substack{i=1,\cdots,N \\ j=1,\cdots,s}}$ とする．次の (1), (2) は同値である．

(1)　$\mathbf{a}_1, \cdots, \mathbf{a}_s$ は線形独立．

(2)　$A$ の $N$ 個の行の中から $s$ 個の行を選んだ行列式

$$\begin{vmatrix} a_{i_1 1} & \cdots & a_{i_1 s} \\ \vdots & \ddots & \vdots \\ a_{i_s 1} & \cdots & a_{i_s s} \end{vmatrix}, \tag{9.3}$$

すなわち $A$ の $s$ 次小行列式の中に 0 でないものが存在する．

**証明．** $s=N$ の場合，(9.3) の行列式は $|A|$ に他ならないので，定理 9.3 より (1) と (2) は同値である．$s<N$ の場合を示す．(1) $\Rightarrow$ (2) を示す．定理 9.8 より，$\boldsymbol{K}^N = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{\nu_1}, \cdots, \mathbf{e}_{\nu_{N-s}}\}$ とできる．したがって，

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{\nu_1}, \cdots, \mathbf{e}_{\nu_{N-s}}) \neq 0$$

である．すなわち

$$0 \neq \begin{vmatrix} a_{11} & \cdots & a_{1s} & & & \\ a_{21} & \cdots & a_{2s} & \mathbf{e}_{\nu_1} & \cdots & \mathbf{e}_{\nu_{N-s}} \\ \vdots & & \vdots & & & \\ a_{N1} & \cdots & a_{Ns} & & & \end{vmatrix}$$

$$= \pm \begin{vmatrix} a_{\nu_1 1} & \cdots & a_{\nu_1 s} & 1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ a_{\nu_{N-s} 1} & \cdots & a_{\nu_{N-s} s} & 0 & & 1 \\ a_{i_1 1} & \cdots & a_{i_1 s} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{i_s 1} & \cdots & a_{i_s s} & 0 & \cdots & 0 \end{vmatrix}$$

$$= \pm \begin{vmatrix} a_{i_1 1} & \cdots & a_{i_1 s} \\ \vdots & \ddots & \vdots \\ a_{i_s 1} & \cdots & a_{i_s s} \end{vmatrix} \tag{9.4}$$

である．次に (2) ⇒ (1) を示す．0 でない小行列式を (9.4) とおく．このとき $1, 2, \cdots, N$ のうち $i_1, \cdots, i_s$ 以外の番号を $\nu_1, \cdots, \nu_{N-s}$ とすると，上の等式より

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{\nu_1}, \cdots, \mathbf{e}_{\nu_{N-s}}) \neq 0$$

となり，$\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{\nu_1}, \cdots, \mathbf{e}_{\nu_{N-s}}$ の線形独立性が得られる．したがって，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ も線形独立である．■

この定理と系 7.7 (3) より次のことがわかる．

**系 9.10**　$1 \leq s \leq N$ とし，$\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_s$ が線形独立であるための必要十分条件は，グラム行列 $G(\mathbf{a}_1, \cdots, \mathbf{a}_s)$ が

$$\det G(\mathbf{a}_1, \cdots, \mathbf{a}_s) > 0$$

をみたすことである．

## 9.3　線形部分空間とその基底

$\boldsymbol{K}^N$ の線形部分空間の定義からはじめる．

**定義 9.1**　$\boldsymbol{K}^N$ の空でない部分集合 $X$ が，任意の $\mathbf{x}, \mathbf{y} \in X$ に対して

$$\mathbf{x} + \mathbf{y} \in X$$

をみたし，さらに任意の $\mathbf{x} \in X$ と任意の $\alpha \in \boldsymbol{K}$ に対して

$$\alpha\,\mathbf{x} \in X$$

をみたすとき，$X$ を $\boldsymbol{K}^N$ の**線形部分空間**あるいは**部分空間**という．

定義より明らかに $\boldsymbol{K}^N$ 自身も $\boldsymbol{K}^N$ の部分空間である．特に $\boldsymbol{K}^N \supsetneq X$ なる部分空間を**真部分空間**という．

$\mathbf{x} \in X$ と $\alpha = 0$ に対して $\alpha\,\mathbf{x} = \mathbf{0}_N$ であるから，部分空間は必ず $\mathbf{0}_N$ を含んでいる．$\{\mathbf{0}_N\}$ は $\boldsymbol{K}^N$ の部分空間であるが，これを**自明な部分空間**という．自明でない部分空間，すなわち $\mathbf{0}_N$ 以外の元を含む部分空間を**非自明な部分空間**という．

**例 9.1**　$l, m, n$ を 0 でない実数とする．

$$X = \left\{ \mathbf{x} : \mathbf{x} \in \boldsymbol{R}^3,\ \frac{x_1}{l} = \frac{x_2}{m} = \frac{x_3}{n} \right\}$$

とする．ただし，$\mathbf{x} = (x_1, x_2, x_3)^T$ と表している．$X$ は $\boldsymbol{R}^3$ の部分空間である ( 注：$X$ は原点 $\mathbf{0}$ を通る直線を表している )．

**証明.** $\mathbf{x} = (x_1, x_2, x_3)^T$ , $\mathbf{y} = (y_1, y_2, y_3)^T \in X$ とする．このとき

$$\frac{x_1}{l} + \frac{y_1}{l} = \frac{x_2}{m} + \frac{y_2}{m} = \frac{x_3}{n} + \frac{y_3}{n}$$

より $\mathbf{x} + \mathbf{y} \in X$ であることがわかる．また $\alpha \in \boldsymbol{R}$ に対して

$$\frac{\alpha x_1}{l} = \frac{\alpha x_2}{m} = \frac{\alpha x_3}{n}$$

であるから，$\alpha\,\mathbf{x} \in X$ が成り立つ．ゆえに $X$ は $\boldsymbol{R}^3$ の部分空間である．■

**例 9.2**　$a, b, c \in \boldsymbol{R}$ を $a = b = c = 0$ ではないものとする．このとき

$$Y = \left\{ \mathbf{x} : \mathbf{x} \in \boldsymbol{R}^3,\ ax_1 + bx_2 + cx_3 = 0 \right\}$$

は $\boldsymbol{R}^3$ の部分空間である ( 注：$Y$ は原点 $\mathbf{0}$ を通る平面を表している )．

**証明.** $\mathbf{x} = (x_1, x_2, x_3)^T$ , $\mathbf{y} = (y_1, y_2, y_3)^T \in X, \alpha \in \boldsymbol{R}$ とする．このとき

$$a\,(x_1 + y_1) + b\,(x_2 + y_2) + c\,(x_3 + y_3) = 0$$
$$a\,(\alpha\,x_1) + b\,(\alpha\,x_2) + c\,(\alpha\,x_3) = 0$$

より $\mathbf{x} + \mathbf{y} \in Y, \alpha\,\mathbf{x} \in Y$ が成り立つ．ゆえに $Y$ は $\boldsymbol{R}^3$ の部分空間である．■

**例 9.3** $\mathbf{a} = (a_1, a_2, a_3)^T \in \boldsymbol{R}^3$ とする．中心 $\mathbf{a}$ で半径 1 の球面

$$X = \left\{\mathbf{x} : \mathbf{x} \in \boldsymbol{R}^3,\ (x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2 = 1\right\}$$

は $\boldsymbol{R}^3$ の部分空間ではない．ただし $\mathbf{x} = (x_1, x_2, x_3)^T$ と表している．

**証明.** $\mathbf{x} = (a_1 + 1, a_2, a_3)^T$ とすると $\mathbf{x} \in X$ である．しかし $2\mathbf{x}$ は

$$(a_1 + 2)^2 + a_2^2 + a_3^2 \geq 4 > 1$$

であるから $X$ には属さない． ■

次の定理は線形包の定義から明らかであろう．

**定理 9.11** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ であるとき，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ は $\boldsymbol{K}^N$ の部分空間である．

**定義 9.2** $\mathbf{a}_1, \cdots, \mathbf{a}_s \in \boldsymbol{K}^N$ のとき，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ を $\mathbf{a}_1, \cdots, \mathbf{a}_s$ **で生成される部分空間**あるいは $\mathbf{a}_1, \cdots, \mathbf{a}_s$ **で張られる部分空間**という．

線形代数で重要な役割を果たす部分空間の例として次のものがある．

**定理 9.12** $A$ を $(M, N)$ 行列とする．このとき定義 3.4 で定めた

$$\mathcal{R}(A) = \left\{A\mathbf{x} : \mathbf{x} \in \boldsymbol{K}^N\right\}$$

は $\boldsymbol{K}^M$ の部分空間である．また

$$\mathcal{N}(A) = \left\{\mathbf{x} : \mathbf{x} \in \boldsymbol{K}^N,\ A\mathbf{x} = \mathbf{0}_M\right\}$$

は $\boldsymbol{K}^N$ の部分空間である．

**証明.** $\mathcal{R}(A) \ni \mathbf{y}, \mathbf{y}'$ ならば，ある $\mathbf{x}, \mathbf{x}' \in \boldsymbol{K}^N$ で $\mathbf{y} = A\mathbf{x}$, $\mathbf{y}' = A\mathbf{x}'$ をみたすものが存在する．したがって，

$$\mathbf{y} + \mathbf{y}' = A\mathbf{x} + A\mathbf{x}' = A(\mathbf{x} + \mathbf{x}')$$

であるから，$\mathbf{y} + \mathbf{y}' \in \mathcal{R}(A)$ である．また $\alpha \in \boldsymbol{K}$ に対して

$$\alpha\,\mathbf{y} = \alpha\,A\mathbf{x} = A(\alpha\,\mathbf{x})$$

より $\alpha\,\mathbf{y} \in \mathcal{R}(A)$ となる．

$\mathcal{N}(A) \ni \mathbf{x}, \mathbf{x}'$ に対して，$A(\mathbf{x} + \mathbf{x}') = A\mathbf{x} + A\mathbf{x}' = \mathbf{0}$ より $\mathbf{x} + \mathbf{x}' \in \mathcal{N}(A)$ である．また，$\alpha \in \boldsymbol{K}$ に対して，$A(\alpha\,\mathbf{x}) = \alpha\,A\mathbf{x} = \mathbf{0}$ より $\alpha\,\mathbf{x} \in \mathcal{N}(A)$ である． ■

**定義 9.3**　$X$ を $\boldsymbol{K}^N$ の部分空間とする．もし，線形独立な $s$ 個のベクトル $\mathbf{u}_1, \cdots, \mathbf{u}_s \in X$ で，

$$\mathrm{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_s\} = X$$

となるものが存在するとき，$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ を $X$ の (一組の) **基底**という．あるいは単に $\mathbf{u}_1, \cdots, \mathbf{u}_s$ を $X$ の (一組の) **基底**という．

**例 9.4**　$\{\mathbf{e}_1, \cdots, \mathbf{e}_N\}$ は $\boldsymbol{K}^N$ の一組の基底である．これを $\boldsymbol{K}^N$ の**標準基底**という．

基底であるための次の必要十分条件は有用なものである．

**定理 9.13**　$X$ を $\boldsymbol{K}^N$ の部分空間とし，$\mathbf{u}_1, \cdots, \mathbf{u}_s \in X$ とする．このとき次の (1), (2) は同値である．

(1)　$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ は $X$ の基底である．

(2)　任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x} = \sum_{j=1}^{s} \alpha_j \mathbf{u}_j$$

をみたす $\alpha_1, \cdots, \alpha_s \in \boldsymbol{K}$ がただ一組存在する．

**証明.** (1) ⇒ (2)．$X = \mathrm{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ より，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，$\mathbf{x} = \sum_{j=1}^{s} \alpha_j \mathbf{u}_j$ をみたす $\alpha_j \in \boldsymbol{K}$ $(j = 1, \cdots, s)$ が存在する．もし $\alpha'_j \in \boldsymbol{K}, j = 1, \cdots, s$ が $\mathbf{x} = \sum_{j=1}^{s} \alpha'_j \mathbf{u}_j$ をみたすとすると，

$$\sum_{j=1}^{s} (\alpha_j - \alpha'_j)\, \mathbf{u}_j = \sum_{j=1}^{s} \alpha_j \mathbf{u}_j - \sum_{j=1}^{s} \alpha'_j \mathbf{u}_j = \mathbf{x} - \mathbf{x} = \mathbf{0}$$

であるが，$\mathbf{u}_1, \cdots, \mathbf{u}_s$ が線形独立であることより，$\alpha_j - \alpha'_j = 0, j = 1, \cdots, s$ である．

(2) ⇒ (1)．明らかに $X = \mathrm{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ である．また

$$\sum_{j=1}^{s} \alpha_j \mathbf{u}_j = \mathbf{0} = \sum_{j=1}^{s} 0\mathbf{u}_j$$

であるから，仮定より $\alpha_j = 0, j = 1, \cdots, s$ となり，線形独立である．■

次に非自明な部分空間は必ず基底をもつことを証明するが，これは部分空間の理論の基本的な定理の一つである．

**定理 9.14**　$\boldsymbol{K}^N$ の非自明な部分空間 $X$ は必ず基底をもつ．

**証明.** $X$ の $\mathbf{0}_N$ でない元を $\mathbf{a}_1$ とする．次の二つの場合がありうる．

**(場合 1a)**　$X$ の元 $\mathbf{a}$ で $\mathbf{a}_1, \mathbf{a}$ が線形独立となるようなものは存在しない．

**(場合 1b)**　$X$ の元 $\mathbf{a}$ で $\mathbf{a}_1, \mathbf{a}$ が線形独立となるようなものが存在する．

場合 1a が起こったときは，$X = \mathrm{span}\{\mathbf{a}_1\}$ である．なぜなら任意の $\mathbf{x} \in X$ に対して，$\mathbf{x}$ と $\mathbf{a}_1$ は線形従属であるから，定理 9.6 より $\mathbf{x} \in \mathrm{span}\{\mathbf{a}_1\}$ である．ゆえに $X \subset \mathrm{span}\{\mathbf{a}_1\}$．また明らかに $\mathrm{span}\{\mathbf{a}_1\} \subset X$ であるから，$X = \mathrm{span}\{\mathbf{a}_1\}$ である．ゆえに $\{\mathbf{a}_1\}$ が $X$ の基底となることがわかり，証明が終了する．

場合 1b が起こった場合は，$\mathbf{a}_1$ と線形独立な $X$ の元の一つを任意にとり，それを $\mathbf{a}_2$ とおく．このとき次の二つの場合が起こりうる．

**(場合 2a)**　$X$ の元 $\mathbf{a}$ で $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}$ が線形独立となるようなものは存在しない．

**(場合 2b)**　$X$ の元 $\mathbf{a}$ で $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}$ が線形独立となるようなものが存在する．

場合 2a が起こったとき，上と同様に定理 9.6 より，$X = \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2\}$ である．したがって，$\{\mathbf{a}_1, \mathbf{a}_2\}$ が $X$ の基底となり，証明が終わる．

場合 2b が起こった場合，$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ が線形独立となるように $\mathbf{a}_3$ をとる．

以下これまでの議論を繰り返す．線形独立なベクトルは $N+1$ 個以上は存在しないので，議論は $N$ 回以下の回数で終わり，定理の証明が終了する．■

このように非自明な部分空間の基底は必ず存在する．しかし，それはただ一組とは限らない．多くの基底が存在しうる．たとえば $X = \mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ は $\boldsymbol{K}^N$ の部分空間であり，明らかに $\{\mathbf{e}_1, \mathbf{e}_2\}$ は $X$ の一組の基底である．しかし，

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

も $X$ の基底である．実際，$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0$ であるから，$\mathbf{a}_1, \mathbf{a}_2$ は線形独立であり，

$$\mathbf{e}_1 = 2\mathbf{a}_1 - \mathbf{a}_2, \quad \mathbf{e}_2 = -\mathbf{a}_1 + \mathbf{a}_2,$$
$$\mathbf{a}_1 = \mathbf{e}_1 + \mathbf{e}_2, \quad \mathbf{a}_2 = \mathbf{e}_1 + 2\mathbf{e}_2$$

より，$X = \mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2\}$ であるから，$\mathbf{a}_1, \mathbf{a}_2$ も $X$ の基底である．

**注意 9.1**　第 18 章で学ぶ離散フーリエ解析，あるいは特に第 19 章で述べる離散ウェーブレット解析では，$\boldsymbol{K}^N$ の標準基底でない基底が中心的な役割を果たす．標準基底以外の基底の構成の方法はいろいろあるが，目的とする対象に適した基底を見出すことは，現在もウェーブレット解析という分野などで研究がすすめられている．

ところで，基底のとり方は一つとは限らないが，基底を構成するベクトルの数は一定である．すなわち次の定理が成り立つ．

**定理 9.15**　$X$ を $\boldsymbol{K}^N$ の非自明な部分空間とする．もし $\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ が $X$ の基底であり，$\{\mathbf{b}_1, \cdots, \mathbf{b}_r\}$ も $X$ の基底であるとき，$s = r$ でなければならない．

**証明.** 仮定より，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\} = \mathrm{span}\{\mathbf{b}_1, \cdots, \mathbf{b}_r\}$ である．まず $r \geq s$ であることを示す．そのため，$r < s$ と仮定して矛盾を導く．$r < s$ の場合 $r < s \leq N$ であるから，定理 9.8 より $\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}$ が線形独立で，かつ

$$\mathrm{span}\{\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}\} = \boldsymbol{K}^N$$

となるように $\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}$ を選べる．明らかに

$$\begin{aligned}\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}\} &= \mathrm{span}\{\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}\} \\ &= \boldsymbol{K}^N\end{aligned}$$

である．$s + N - r > s + N - s = N$ であるから，$\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}$ は線形従属である．そこで，$\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t}$ が線形従属になるような最初の番号 $i_t$ を選ぶ．$i_t \geq i_1$ である．$i_t > i_1$ の場合は

$$\mathbf{e}_{i_t} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{t-1}}\}$$

であるから，

$$\mathbf{e}_{i_t} \in \mathrm{span}\{\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_t-1}\}$$

となり，$\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}$ が線形独立であることに反する．また $i_t = i_1$ の場合は $\mathbf{e}_{i_t} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ であるから，$\mathbf{e}_{i_t} \in \mathrm{span}\{\mathbf{b}_1, \cdots, \mathbf{b}_s\}$ となる．この場合も $\mathbf{b}_1, \cdots, \mathbf{b}_r, \mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_{N-r}}$ が線形独立であることに反する．ゆえに

$r \geq s$ となることが証明された．

次に $\mathbf{a}_1, \cdots, \mathbf{a}_s$ と $\mathbf{b}_1, \cdots, \mathbf{b}_r$ の役割を交換して同様に議論すれば $r \leq s$ も示せる．よって $s = r$ である． ■

この事実に基づいて，部分空間に次元という量が次のように定義される．

**定義 9.4** $X$ を $\boldsymbol{K}^N$ の非自明な部分空間とする．$X$ が基底 $\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ をもつとき，$X$ の**次元**は $s$ であるといい，

$$\dim X = s$$

と表す．なお便宜上，自明な部分空間 $\{\mathbf{0}\}$ の次元は 0 と定めておく．

**定理 9.16** $X, Y \subset \boldsymbol{K}^N$ を部分空間で $X \subset Y$ とする．このとき，$\dim X \leq \dim Y$ である．また $X \subset Y$ で，$\dim X = \dim Y$ ならば $X = Y$ である．

**証明.** $X$ が自明な部分空間の場合は明らかであるから，非自明な場合を証明する．$X$ の基底 $\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ をとる．$Y$ に対して定理 9.14 の証明の議論を当てはめる．$\mathbf{a}_1, \cdots, \mathbf{a}_s$ は $Y$ の線形独立な要素であるから，定理 9.14 の証明の記号を用いれば，$Y$ に関して場合 sa と場合 sb が起こる．したがって，そこでの議論から，$\mathbf{a}_1, \cdots, \mathbf{a}_s$ を含む $s$ 個以上のベクトルからなる $Y$ の基底をとれることがわかる．ゆえに $\dim X \leq \dim Y$ である．

次に後半の主張を示す．$\dim X = \dim Y$ とする．$X \subsetneq Y$ であるとすると，ある $\mathbf{y} \in Y$ が存在して，$\mathbf{y} \notin \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ である．したがって，定理 9.6 より $\mathbf{y}, \mathbf{a}_1, \cdots, \mathbf{a}_s$ は $Y$ の線形独立な元である．よって

$$\dim X + 1 = \dim \mathrm{span}\{\mathbf{y}, \mathbf{a}_1, \cdots, \mathbf{a}_s\} \leq \dim Y = \dim X$$

となり，矛盾が導かれる． ■

本書を通して次の記号を用いる．集合 $X$ とその部分集合 $Y$ に対して，

$$X \setminus Y = \{x : x \in X \text{ かつ } x \notin Y\} \tag{9.5}$$

とする．

**定理 9.17** $\mathbf{a}_1, \cdots, \mathbf{a}_r \in \boldsymbol{K}^N \setminus \{\mathbf{0}\}$ とする．$s = \dim \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ とする．もし $s < r$ ならば，$s$ 個の $\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(s)}$ を選んで，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ の基底とすることができる．

**証明.** $X = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_r$ は線形従属である．なぜならば，もし線形独立ならば $\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ が基底になり，$s = r$ となってしまう．そこで，$\mathbf{a}_1, \cdots, \mathbf{a}_t$ が線形従属となる最初の番号 $t$ を $t_1$ とする．明らかに $t_1 \geq 2$ である．$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}$ は線形独立で，$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1}$ が線形従属であるから，$\mathbf{a}_{t_1} \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}\}$ である．したがって，

$$X = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_r\}$$

である．もし $\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_r$ が線形独立ならば，これが $X$ の基底になり，$r - 1 = s$ である．もし線形従属ならば，$\mathbf{a}_1, \cdots, \mathbf{a}_{t_1-1}, \mathbf{a}_{t_1+1}, \cdots, \mathbf{a}_t$ が線形従属になる最初の番号を $t_2$ とする．$t_2 \geq t_1 + 1$ である．前と同様の議論で，

$$X = \mathrm{span}\ (\{\mathbf{a}_1, \cdots, \mathbf{a}_r\} \setminus \{\mathbf{a}_{t_1}, \mathbf{a}_{t_2}\})$$

となることが示せる．もし $\mathbf{a}_1, \cdots, \mathbf{a}_r$ から $\mathbf{a}_{t_1}, \mathbf{a}_{t_2}$ を除いたベクトルが線形独立ならば，それが $X$ の基底になる．そうでなければ，上の議論を繰り返していけば，有限回のうちに線形独立でかつ $X$ を生成するようなものを $\mathbf{a}_1, \cdots, \mathbf{a}_r$ の中から抜き出すことができる．定義よりそれが $X$ の基底となっている．■

部分空間の基底に関して次の定理 9.18，定理 9.19 はしばしば線形代数の理論的な考察に用いられる．

**定理 9.18**　$X$ を $\boldsymbol{K}^N$ の部分空間で，$s = \dim X \geq 1$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_s \in X$ が線形独立ならば，$\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ は $X$ の基底になっている．

**証明.** 明らかに $X \supset \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ である．もし $X \underset{\neq}{\supset} \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ であるとすると，ある $\mathbf{y} \in X$ が存在し $\mathbf{y} \notin \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ であるから，定理 9.6 より $\mathbf{y}, \mathbf{a}_1, \cdots, \mathbf{a}_s$ は線形独立になる．このことは $\dim X \geq s+1$ となることを示しているので $s = \dim X$ という仮定に反する．ゆえに $X = \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ でなければならない．■

**定理 9.19**　$X$ を $\boldsymbol{K}^N$ の部分空間で，$s = \dim X \geq 1$ とする．$\mathbf{a}_1, \cdots, \mathbf{a}_r \in X$ (ただし $r < s$) が線形独立ならば，適当な $\mathbf{a}_{r+1}, \cdots, \mathbf{a}_s \in X$ を選んで $\{\mathbf{a}_1, \cdots, \mathbf{a}_s\}$ が $X$ の基底になるようにすることができる．

**証明.** 定理 9.16 の前半の証明および定理 9.15 より明らか．■

二つの部分空間の直和を定義する．直和は重要な概念である．

**定義 9.5**　$X, X_0, X_1 \subset \boldsymbol{K}^N$ を部分空間で，$X_0, X_1 \subset X$ とする．もし

(i)　任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，ある $\mathbf{x}' \in X_0$, $\mathbf{x}'' \in X_1$ が存在し，$\mathbf{x} = \mathbf{x}' + \mathbf{x}''$,

(ii)　$X_0 \cap X_1 = \{\mathbf{0}\}$

が成り立つとき，$X$ は $X_0$ と $X_1$ の**直和**である，あるいは $X$ は $X_0$ と $X_1$ に**直和分解**されるといい，

$$X = X_0 \dotplus X_1$$

と表す．

直和分解に関していくつかの基本的な性質を述べておく．

**命題 9.20**　$X, X_0, X_1 \subset \boldsymbol{K}^N$ を部分空間で，$X_0, X_1 \subset X$ とする．$X = X_0 \dotplus X_1$ であるための必要十分条件は，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，$\mathbf{x}' \in X_0$, $\mathbf{x}'' \in X_1$ で，$\mathbf{x} = \mathbf{x}' + \mathbf{x}''$ となるものが一意的に存在することである．

**証明.**　$X = X_0 \dotplus X_1$ であるとする．もし $\mathbf{x} = \mathbf{x}' + \mathbf{x}'' = \mathbf{y}' + \mathbf{y}''$ ($\mathbf{x}', \mathbf{y}' \in X_0$, $\mathbf{x}'', \mathbf{y}'' \in X_1$) と表せたとする．このとき，$\mathbf{x}' - \mathbf{y}' \in X_0$ かつ $\mathbf{x}' - \mathbf{y}' = \mathbf{y}'' - \mathbf{x}'' \in X_1$ であるから，$\mathbf{x}' - \mathbf{y}' = \mathbf{0}$ である．これより $\mathbf{x}' = \mathbf{y}'$, $\mathbf{x}'' = \mathbf{y}''$ が得られる．逆を示す．もし $\mathbf{x} \in X_0 \cap X_1$ とすると，$\mathbf{x} = \mathbf{x} + \mathbf{0}$ ($\mathbf{x} \in X_0$, $\mathbf{0} \in X_1$) であり，$\mathbf{x} = \mathbf{0} + \mathbf{x}$ ($\mathbf{0} \in X_0$, $\mathbf{x} \in X_1$) でもあるから，一意性より $\mathbf{x} = \mathbf{0}$ が得られる．ゆえに $X_0 \cap X_1 = \{\mathbf{0}\}$ となり，$X = X_0 \dotplus X_1$ となる．■

**系 9.21**　$X, X_0, X_1 \subset \boldsymbol{K}^N$ を部分空間で，$X = X_0 \dotplus X_1$ とする．このとき

$$\dim X = \dim X_0 + \dim X_1$$

である．

**証明.** $\dim X_0 = r$, $\dim X_1 = s$ とする．$\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ を $X_0$ の基底, $\{\mathbf{v}_1, \cdots, \mathbf{v}_s\}$ を $X_1$ の基底とする．いま

$$\sum_{j=1}^{r} c_j \mathbf{u}_j + \sum_{k=1}^{s} d_k \mathbf{v}_k = \mathbf{0}$$

が成立しているとする．このとき，$\displaystyle\sum_{j=1}^{r} c_j \mathbf{u}_j = -\sum_{k=1}^{s} d_k \mathbf{v}_k \in X_1$ でもあるから，$\displaystyle\sum_{j=1}^{r} c_j \mathbf{u}_j = \mathbf{0}$ である．したがって $\displaystyle\sum_{k=1}^{s} d_k \mathbf{v}_k = \mathbf{0}$ も成り立つ．ゆえに $c_1 = \cdots = c_r = 0$, $d_1 = \cdots = d_s = 0$ となり，

$$\mathbf{u}_1, \cdots, \mathbf{u}_r, \quad \mathbf{v}_1, \cdots, \mathbf{v}_s \tag{9.6}$$

は線形独立である．また明らかにこれらのベクトルの線形包は $X = X_0 \dot{+} X_1$ と一致している．ゆえに (9.6) は $X$ の基底であることがわかる．よって

$$\dim X = r + s = \dim X_0 + \dim X_1$$

である．■

最後に部分空間上の線形写像を定義しておこう．

**定義 9.6**　$X$ を $\boldsymbol{K}^N$ の部分空間とする．$X$ から $\boldsymbol{K}^M$ への写像 $T$ が線形写像であるとは，任意の $\mathbf{x}, \mathbf{x}' \in X$, および任意の $\alpha \in \boldsymbol{K}$ に対して

$$T(\mathbf{x} + \mathbf{x}') = T(\mathbf{x}) + T(\mathbf{x}'),$$
$$T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$$

をみたすことである．

たとえば次のような写像を考えてみよう．

**定義 9.7**　$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への写像とする．$X$ を $\boldsymbol{K}^N$ の部分集合とする．このとき，$X$ の元 $\mathbf{x}$ に $T(\mathbf{x})$ を対応させる $X$ から $\boldsymbol{K}^M$ への写像を $T|_X$ と表し，$T$ の $X$ への**制限**という．

$T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像とし，$X$ を $\boldsymbol{K}^N$ の部分空間とする．このとき $T|_X$ は $X$ から $\boldsymbol{K}^N$ への線形写像である．

# 第10章　内積と正規直交基底

この章では内積とそれにかかわる線形代数の基礎事項を述べる．内積は応用上もきわめて重要な概念である．

**ポイント10**　直交性にまつわる話題は有用なものが多い．具体的な応用については第2部でも示してあるが，特に第3部に詳しく記されている．基底と正規直交基底の違いが説明できるようになってほしい．

## 10.1　内積と直交性

平面幾何では，ベクトル $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ の長さが $|\overrightarrow{x}| = \sqrt{|x_1|^2 + |x_2|^2}$ により定義されている．また，二つのベクトル $\overrightarrow{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $\overrightarrow{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ に対して，その内積が

$$\overrightarrow{a} \cdot \overrightarrow{b} = |\overrightarrow{a}| \left|\overrightarrow{b}\right| \cos\theta \tag{10.1}$$

により定義されている．ここで $\theta$ は $\overrightarrow{a}$ と $\overrightarrow{b}$ のなす角で，$0 \leq \theta \leq \pi$ をみたすものである．高等学校で学んだように

$$\overrightarrow{a} \cdot \overrightarrow{b} = a_1 b_1 + a_2 b_2 \tag{10.2}$$

と表されている．

$\boldsymbol{K}^N$ に対しても，(10.2) に倣って，次のように内積という量が定義される．

**定義 10.1**　$\mathbf{x} = (x_j)_{j=1,\cdots,N}$ , $\mathbf{y} = (y_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ に対して，

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^N} = \sum_{j=1}^{N} x_j \, \overline{y_j}$$

を $\mathbf{x}$ と $\mathbf{y}$ の**内積**，あるいは**ユークリッド内積**という．特に混乱が生じない場合は $\langle \mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^N}$ を $\langle \mathbf{x}, \mathbf{y} \rangle$ と略記する．

また，

$$\|\mathbf{x}\| = \sqrt{\sum_{j=1}^{N} |x_j|^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

を $\mathbf{x}$ の**長さ**あるいは $l^2$ **ノルム**，$\langle \mathbf{x}, \mathbf{x} \rangle$ を $\mathbf{x}$ の**エネルギー**という．

内積の定義から次のことがわかる．

**命題 10.1**　(1)　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して

$$\mathbf{x}^* \, \mathbf{y} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = \langle \mathbf{y}, \mathbf{x} \rangle .$$

特に $\mathbf{x}, \mathbf{y} \in \boldsymbol{R}^N$ ならば

$$\mathbf{x}^T \, \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle .$$

(2)　任意の $\mathbf{y} \in \boldsymbol{K}^N$ に対して $\langle \mathbf{0}, \mathbf{y} \rangle = 0$ である．また $\mathbf{x} \in \boldsymbol{K}^N$ が任意の $\mathbf{y} \in \boldsymbol{K}^N$ に対して $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ をみたすならば，$\mathbf{x} = \mathbf{0}$ である．

(3)　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$, $\alpha, \beta \in \boldsymbol{K}$ に対して

$$\langle \alpha \mathbf{x}, \, \beta \mathbf{y} \rangle = \alpha \overline{\beta} \, \langle \mathbf{x}, \, \mathbf{y} \rangle ,$$
$$\|\alpha \, \mathbf{x}\| = |\alpha| \, \|\mathbf{x}\| .$$

**証明.** (1) $\mathbf{x} = (x_j)_{j=1,\cdots,N}$ , $\mathbf{y} = (y_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ とすると，

$$\mathbf{x}^* \, \mathbf{y} = \sum_{j=1}^{N} \overline{x_j} y_j = \overline{\sum_{j=1}^{N} x_j \overline{y_j}} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle},$$
$$\sum_{j=1}^{N} \overline{x_j} y_j = \sum_{j=1}^{N} y_j \overline{x_j} = \langle \mathbf{y}, \mathbf{x} \rangle$$

が成り立っている．(2) の前半の主張は明らかであるから，(2) の後半の主張を証明する．$\mathbf{x} = (x_j)_{j=1,\cdots,N}$ とすると，基本ベクトル $\mathbf{e}_j^{(N)}$ に対して

$$x_j = \left\langle \mathbf{x}, \mathbf{e}_j^{(N)} \right\rangle = 0, \quad j = 1, \cdots, N$$

となってしまう．よって $\mathbf{x} = \mathbf{0}$ である．(3) は内積の定義から明らかである．■

平面幾何の場合，二つのベクトルの内積とそのなす角 $\theta \in [0, \pi]$ の間には

$$\overrightarrow{a} \cdot \overrightarrow{b} = |\overrightarrow{a}| \left|\overrightarrow{b}\right| \cos\theta$$

の関係があることは高校で学んだ．したがって，長さが 0 でないベクトル $\overrightarrow{a}$ と $\overrightarrow{b}$ に対して，

$$\overrightarrow{a} \cdot \overrightarrow{b} = 0 \Longleftrightarrow \theta = \frac{\pi}{2} \Longleftrightarrow \overrightarrow{a} \text{と} \overrightarrow{b} \text{は直交}$$

となっている．そこで一般の場合にもベクトルのなす角と直交性を定義するが，その前に次の重要な不等式を証明しておく．

**定理 10.2**（シュヴァルツの不等式） $\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| . \tag{10.3}$$

ただし，$\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ で等号が成り立つのは $\mathbf{x}$ と $\mathbf{y}$ が線形従属の場合に限る．

**証明.**

$$\begin{aligned}\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 &= \det \begin{pmatrix} \|\mathbf{x}\|^2 & \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\ \langle \mathbf{x}, \mathbf{y} \rangle & \|\mathbf{y}\|^2 \end{pmatrix} = \det \begin{pmatrix} \langle \mathbf{x}, \mathbf{x} \rangle & \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\ \overline{\langle \mathbf{y}, \mathbf{x} \rangle} & \langle \mathbf{y}, \mathbf{y} \rangle \end{pmatrix} \\ &= \det G(\mathbf{x}, \mathbf{y})\end{aligned}$$

であるが，系 7.7 より $\det G(\mathbf{x}, \mathbf{y}) \geq 0$ が成り立っている．したがって (10.3) が証明された．また系 9.10 より等号条件も得られる． ■

**定義 10.2** $\mathbf{x}, \mathbf{y} \in \boldsymbol{R}^N$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ に対して，

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}, \quad 0 \leq \theta \leq \pi$$

をみたす $\theta$ がただ一つ存在するが，この $\theta$ を $\mathbf{x}$ と $\mathbf{y}$ のなす角といい，

$$\angle (\mathbf{x}, \mathbf{y})$$

で表す．特に $\angle (\mathbf{x}, \mathbf{y}) = \dfrac{\pi}{2}$ の場合，すなわち $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ をみたすとき，$\mathbf{x}$ と $\mathbf{y}$ は**直交する**という．

**定義 10.3** $\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ の場合，

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0$$

をみたすことを $\mathbf{x}$ と $\mathbf{y}$ が**直交する**と定義し，$\mathbf{x} \perp \mathbf{y}$ と表す．

**例 10.1** $\boldsymbol{K}^N$ の基本ベクトル $\mathbf{e}_i^{(N)}$, $\mathbf{e}_j^{(N)}$ は $i \neq j$ のとき直交している．

**例 10.2**　$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ と $\mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ のなす角は,

$$\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|} = \frac{1}{\sqrt{2}} = \cos \frac{\pi}{4}$$

より $\dfrac{\pi}{4}$ である.

**例 10.3**　$\mathbf{x} = \begin{pmatrix} a \\ \overline{b} \\ c \end{pmatrix}, \mathbf{y} = \begin{pmatrix} -b \\ \overline{a} \\ 0 \end{pmatrix} \in \boldsymbol{K}^3$ は直交している.

平面幾何でよく知られた定理にピタゴラスの定理がある．これに相当する定理が次のような形で成り立つ.

**定理 10.3（ピタゴラスの定理）**　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ とし，$\mathbf{z} = \mathbf{x} + \mathbf{y}$ とする．$\mathbf{x}$ と $\mathbf{y}$ が直交しているならば,

$$\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

が成り立つ.

**証明.** ノルムの定義から,

$$\begin{aligned}
\|\mathbf{z}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2 \operatorname{Re} \langle \mathbf{x}, \mathbf{y} \rangle
\end{aligned}$$

が成り立つ．したがって，定理が得られる．■

この定理の証明から次の不等式が得られる.

**定理 10.4（三角不等式）**　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して,

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| . \tag{10.4}$$

また，$\mathbf{x} \neq \mathbf{0}, \mathbf{y} \neq \mathbf{0}$ で等号が成り立つのは，ある $\alpha \geq 0$ が存在し $\mathbf{x} = \alpha \, \mathbf{y}$ が成り立つ場合に限る.

**証明.** 定理 10.3 の証明とシュヴァルツの不等式より

$$\|\mathbf{x}+\mathbf{y}\|^2 \le \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\,|\langle \mathbf{x},\mathbf{y}\rangle| \le \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\,\|\mathbf{x}\|\,\|\mathbf{y}\|$$
$$= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$$

である．いま (10.4) において等号が成り立つとすると，$|\langle \mathbf{x},\mathbf{y}\rangle| = \|\mathbf{x}\|\,\|\mathbf{y}\|$ であるから，定理 10.2 の等号条件から $\mathbf{x} = \alpha\,\mathbf{y}$ をみたす $\alpha \in \boldsymbol{K}$ が存在する．このとき，

$$(|\alpha|+1)\,\|\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\| = \|\mathbf{x}+\mathbf{y}\| = |\alpha+1|\,\|\mathbf{y}\|$$

より，$|\alpha|+1 = |\alpha+1|$ である．よって $\alpha \ge 0$ でなければならない．逆に，$\mathbf{x} = \alpha\,\mathbf{y}$, $\alpha \ge 0$ となっている場合は

$$\|\mathbf{x}+\mathbf{y}\| = \|(\alpha+1)\,\mathbf{y}\| = (\alpha+1)\,\|\mathbf{y}\| = \|\alpha\,\mathbf{y}\| + \|\mathbf{y}\|$$
$$= \|\mathbf{x}\| + \|\mathbf{y}\|$$

が成り立っている．■

**系 10.5** $\mathbf{x}^{(j)} \in \boldsymbol{K}^N$, $j = 1, \cdots, s$ に対して，

$$\left\|\sum_{j=1}^{s} \mathbf{x}^{(j)}\right\| \le \sum_{j=1}^{s} \left\|\mathbf{x}^{(j)}\right\|.$$

**証明.** $s=1$ の場合は明らか．$s-1$ の場合に正しいとして，$s$ の場合を証明する．定理 10.4 より

$$\left\|\sum_{j=1}^{s} \mathbf{x}^{(j)}\right\| = \left\|\sum_{j=1}^{s-1} \mathbf{x}^{(j)} + \mathbf{x}^{(s)}\right\| \le \left\|\sum_{j=1}^{s-1} \mathbf{x}^{(j)}\right\| + \left\|\mathbf{x}^{(s)}\right\|$$
$$\le \sum_{j=1}^{s-1} \left\|\mathbf{x}^{(j)}\right\| + \left\|\mathbf{x}^{(s)}\right\|.$$

■

内積と共役行列の次の関係式は，今後，頻繁に使う．$A$ を $(M,N)$ 行列とし，$A^*$ を $A$ の共役行列とする ((2.4) 参照)．$A^*$ は $(N,M)$ 行列である．次のことが成り立つ．

**定理 10.6** $A$ を $(M,N)$ 行列とする．任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{y} \in \boldsymbol{K}^M$ に対して，

$$\langle A\mathbf{x}, \mathbf{y}\rangle_{\boldsymbol{K}^M} = \langle \mathbf{x}, A^*\mathbf{y}\rangle_{\boldsymbol{K}^N}.$$

また $B$ が $(N,M)$ 行列のとき，

$$\langle \mathbf{x}, B\mathbf{y} \rangle_{\boldsymbol{K}^N} = \langle B^*\mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^M}.$$

**証明.** $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix}$, $A = (a_{ij})$, $A^* = (a^*_{ij})$ とする．$a^*_{ij} = \overline{a_{ji}}$ である．したがって，

$$\begin{aligned}
\langle A\mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^M} &= \sum_{m=1}^{M} \left( \sum_{n=1}^{N} a_{mn} x_n \right) \overline{y_m} = \sum_{n=1}^{N} x_n \left( \sum_{m=1}^{M} a_{mn} \overline{y_m} \right) \\
&= \sum_{n=1}^{N} x_n \overline{\left( \sum_{m=1}^{M} \overline{a_{mn}} y_m \right)} = \sum_{n=1}^{N} x_n \overline{\left( \sum_{m=1}^{M} a^*_{nm} y_m \right)} \\
&= \langle \mathbf{x}, A^*\mathbf{y} \rangle_{\boldsymbol{K}^N}.
\end{aligned}$$

後半の主張は，$B = (B^*)^*$ と前半の主張より

$$\langle \mathbf{x}, B\mathbf{y} \rangle_{\boldsymbol{K}^N} = \langle \mathbf{x}, (B^*)^* \mathbf{y} \rangle_{\boldsymbol{K}^N} = \langle B^*\mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^M}$$

として得られる．■

[**補足 1**]

$\|\cdot\|$ を $l^2$ ノルムといっているが，じつは一般にノルムという抽象的な概念がある．

**定義 10.4**　$\boldsymbol{K}^N$ から $\boldsymbol{R}$ への関数 $\rho$ が次の条件 (1), (2), (3), (4) をみたすとき，$\boldsymbol{K}^N$ 上の**ノルム**であるという．

(1)　$\rho(\mathbf{x}) \geq 0, \quad \mathbf{x} \in \boldsymbol{K}^N.$

(2)　$\rho(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}.$

(3)　$\rho(a\,\mathbf{x}) = |a|\,\rho(\mathbf{x}), \quad a \in \boldsymbol{K},\ \mathbf{x} \in \boldsymbol{K}^N.$

(4)　$\rho(\mathbf{x} + \mathbf{y}) \leq \rho(\mathbf{x}) + \rho(\mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N.$

すでに示したように $l^2$ ノルムはノルムである．このほかに次のようなノルムもある．

**例 10.4**　$\mathbf{x} = (x_j) \in \boldsymbol{K}^N$ に対して，

$$\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq N} |x_j|,$$

$$\|\mathbf{x}\|_p = \left( \sum_{j=1}^{N} |x_j|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

$\|\cdot\|_p$ を $l^p$ **ノルム**という $(1 \leq p \leq \infty)$.

[**補足 2**]

$\mathbf{x} = (x_j)$, $\mathbf{y} = (y_j) \in \boldsymbol{R}^4$ に対して,

$$\langle \mathbf{x}, \mathbf{y} \rangle_L = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$$

と定義したものをローレンツ計量という．ローレンツ計量は特殊相対性理論で用いられる．特殊相対性理論では，$x_1, x_2, x_3$ は空間の座標を表し，$x_4$ は時間の座標を表している．

**問題 10.1**　$\rho$ を $\boldsymbol{K}^N$ 上のノルムとする．このとき，$\rho$ は $\boldsymbol{K}^N$ 上の連続関数，すなわち，任意の $\varepsilon > 0$ に対して，ある $\delta > 0$ が存在し，

$$\|\mathbf{h}\| < \delta \Longrightarrow |\rho(\mathbf{x} + \mathbf{h}) - \rho(\mathbf{x})| < \varepsilon$$

をみたすことを証明せよ．

**問題 10.2**　$\rho$ を $\boldsymbol{K}^N$ 上のノルムとする．このときある正定数 $c$, $C$ が存在し，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$c\|\mathbf{x}\| \leq \rho(\mathbf{x}) \leq C\|\mathbf{x}\|$$

が成り立つことを証明せよ．

## 10.2　正規直交基底

まず正規直交基底の定義からはじめる．

**定義 10.5**　$X$ を $\boldsymbol{K}^N$ の部分空間とする．$X \ni \mathbf{u}_1, \cdots, \mathbf{u}_s$ が

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

をみたすとき，$X$ の**正規直交系**という．$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ が $X$ の基底でありかつ正規直交系であるとき，$X$ の**正規直交基底**という．

$X \ni \mathbf{u}_1, \cdots, \mathbf{u}_s$ が

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0, \quad i \neq j$$

をみたすとき，$X$ の**直交系**という．また $X$ の基底で直交系をなすものを**直交基**

底という．

**例 10.5**　$\boldsymbol{K}^N$ の基本ベクトル $\left\{\mathbf{e}_1^{(N)},\cdots,\mathbf{e}_N^{(N)}\right\}$ は $\boldsymbol{K}^N$ の正規直交基底である．

**命題 10.7**　$X$ を $\boldsymbol{K}^N$ の部分空間とし，$\mathbf{u}_1,\cdots,\mathbf{u}_s$ を $X$ の直交系で，$\mathbf{u}_j \neq \mathbf{0}, j=1,\cdots,s$ とする．このとき，$\mathbf{u}_1,\cdots,\mathbf{u}_s$ は線形独立である．

特に $s=\dim X$ で，$\mathbf{u}_1,\cdots,\mathbf{u}_s$ が $X$ の正規直交系ならば $X$ の正規直交基底である．

**証明.** $\sum_{j=1}^{s}\lambda_j\mathbf{u}_j=\mathbf{0}$ であるとする．このとき，$k=1,\cdots,s$ に対して

$$0=\left\langle\sum_{j=1}^{s}\lambda_j\mathbf{u}_j,\mathbf{u}_k\right\rangle=\lambda_k\langle\mathbf{u}_k,\mathbf{u}_k\rangle=\lambda_k\|\mathbf{u}_k\|^2.$$

したがって，$\lambda_k=0$．ゆえに $\mathbf{u}_1,\cdots,\mathbf{u}_s$ は線形独立である．$s=\dim X$ ならば，定理 9.18 より $\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ は $X$ の基底でもあるから，直交基底になっている．

■

$X$ を $\boldsymbol{K}^N$ の $s$ 次元部分空間 $(s\geq 1)$ とし，$\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ が正規直交基底であるとする．基底であるから，任意の $\mathbf{x}\in X$ に対して，$\mathbf{x}=\sum_{j=1}^{s}c_j\mathbf{u}_j$ をみたすスカラー $c_1,\cdots,c_s$ が一意的に存在する．一般に $c_j$ を計算することは簡単ではないが，正規直交基底の場合，次のように簡単に計算できる．

**定理 10.8（フーリエの原理）**　$X$ を $\boldsymbol{K}^N$ の $s$ 次元部分空間 $(s\geq 1)$ とし，$\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ が正規直交基底とする．このとき任意の $\mathbf{x}\in X$ に対して，

$$\mathbf{x}=\sum_{j=1}^{s}\langle\mathbf{x},\mathbf{u}_j\rangle\ \mathbf{u}_j$$

が成り立つ．

**証明.** $\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ が基底であることより，$\mathbf{x}=\sum_{j=1}^{s}c_j\mathbf{u}_j$ と表せる．$\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ が正規直交系であることより，

$$\langle\mathbf{x},\mathbf{u}_k\rangle=\left\langle\sum_{j=1}^{s}c_j\mathbf{u}_j,\mathbf{u}_k\right\rangle=\sum_{j=1}^{s}c_j\langle\mathbf{u}_j,\mathbf{u}_k\rangle=c_k$$

が得られる．■

定理 10.8 は，フーリエ解析学でフーリエの反転公式と呼ばれているものに相当

する．後に解説する離散フーリエ解析の基礎となる定理である．

## 10.3 シュミットの直交化法

正規直交基底を構成する方法にはいろいろなものがある．本節では，シュミットの直交化法を述べる．これは $\boldsymbol{K}^N$ の部分空間 $X$ に基底 $\mathbf{w}_1, \cdots, \mathbf{w}_s$ が与えられたとき，それから $X$ の正規直交基底を作る方法である．シュミットの直交化法は非常に強力な手法である．

シュミットの直交化法の基本的な方針は，

$$
\begin{aligned}
\mathbf{v}_1 &= r_{11}\mathbf{w}_1 \\
\mathbf{v}_2 &= r_{21}\mathbf{w}_1 + r_{22}\mathbf{w}_2 \\
&\ \ \vdots \\
\mathbf{v}_s &= r_{s1}\mathbf{w}_1 + \cdots + r_{ss}\mathbf{w}_s
\end{aligned}
\tag{10.5}
$$

が $\mathbf{v}_j \neq 0\,(j = 1, \cdots, s)$ かつ $\langle \mathbf{v}_j, \mathbf{v}_k \rangle = 0\,(j \neq k)$ となるように係数 $r_{mn}$ を選び，$\mathbf{u}_j = \dfrac{\mathbf{v}_j}{\|\mathbf{v}_j\|}$ として $X$ の正規直交基底を作る方法である．

まずはじめに

$$\mathbf{v}_1 = \mathbf{w}_1$$

とおく (すなわち $r_{11} = 1$)．次に

$$\mathbf{v}_2 = r_{21}\mathbf{v}_1 + r_{22}\mathbf{w}_2$$

で，$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ となる $r_{21}$, $r_{22}$ を見出す．

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \overline{r_{21}}\,\|\mathbf{v}_1\|^2 + \overline{r_{22}}\,\langle \mathbf{v}_1, \mathbf{w}_2 \rangle$$

であるから，$r_{22} = 1, \overline{r_{21}} = -\dfrac{\langle \mathbf{v}_1, \mathbf{w}_2 \rangle}{\|\mathbf{v}_1\|^2}$ であれば，$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ をみたす．ゆえに

$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

とおけば，これが求めるベクトルである．また，仮定より $\mathbf{w}_2$ と $\mathbf{w}_1\,(= \mathbf{v}_1)$ は線形独立であるから，$\mathbf{v}_2 \neq \mathbf{0}$ である．次に

$$\mathbf{v}_3 = r_{31}\mathbf{v}_1 + r_{32}\mathbf{v}_2 + r_{33}\mathbf{w}_3$$

とおき，$\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$ となるように係数を選ぶ．

$$\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \overline{r_{31}} \|\mathbf{v}_1\|^2 + \overline{r_{33}} \langle \mathbf{v}_1, \mathbf{w}_3 \rangle,$$
$$\langle \mathbf{v}_2, \mathbf{v}_3 \rangle = \overline{r_{32}} \|\mathbf{v}_2\|^2 + \overline{r_{33}} \langle \mathbf{v}_2, \mathbf{w}_3 \rangle.$$

したがって，$r_{33} = 1$, $\overline{r_{31}} = -\dfrac{\langle \mathbf{v}_1, \mathbf{w}_3 \rangle}{\|\mathbf{v}_1\|^2}$, $\overline{r_{32}} = -\dfrac{\langle \mathbf{v}_2, \mathbf{w}_3 \rangle}{\|\mathbf{v}_2\|^2}$ とすれば，$\langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$ がみたされる．ゆえに

$$\mathbf{v}_3 = \mathbf{w}_3 - \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|} \mathbf{v}_2 - \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|} \mathbf{v}_1$$

が求めるベクトルである．$\mathbf{v}_2, \mathbf{v}_1 \in \mathrm{span}\{\mathbf{w}_1, \mathbf{w}_2\}$ で，仮定より $\mathbf{w}_3$, $\mathbf{w}_2$, $\mathbf{w}_1$ は線形独立であるから，$\mathbf{v}_3 \neq \mathbf{0}$ である．以下，同様にして，$\mathbf{v}_1, \cdots, \mathbf{v}_n$ を

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \ (i \neq j), \quad \mathbf{v}_j \in \mathrm{span}\{\mathbf{w}_1, \cdots, \mathbf{w}_j\}$$

$i, j \in \{1, \cdots, n\}$ となるように構成できれば，

$$\mathbf{v}_{n+1} = \mathbf{w}_{n+1} - \frac{\langle \mathbf{w}_{n+1}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|} \mathbf{v}_n - \frac{\langle \mathbf{w}_{n+1}, \mathbf{v}_{n-1} \rangle}{\|\mathbf{v}_{n-1}\|} \mathbf{v}_{n-1} - \cdots - \frac{\langle \mathbf{w}_{n+1}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|} \mathbf{v}_1$$

により，$\langle \mathbf{v}_j, \mathbf{v}_{n+1} \rangle = 0$, $j = 1, \cdots, n$ となるようにできる．また仮定から $\mathbf{v}_j \in \mathrm{span}\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$, $j = 1, \cdots, n$ であるから，

$$\sum_{j=1}^{n} \frac{\langle \mathbf{w}_{n+1}, \mathbf{v}_j \rangle}{\|\mathbf{v}_j\|} \mathbf{v}_j \in \mathrm{span}\{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$$

であり，$\mathbf{v}_{n+1} \in \mathrm{span}\{\mathbf{w}_1, \cdots, \mathbf{w}_{n+1}\}$ が成り立つ．仮定より $\mathbf{w}_{n+1}, \mathbf{w}_n \cdots, \mathbf{w}_1$ は線形独立であるから，定義より $\mathbf{v}_{n+1} \neq \mathbf{0}$ である．

このようにして，$X$ の直交系

$$\mathbf{v}_1, \cdots, \mathbf{v}_s$$

が得られる．$s = \dim X$ であるから，これは $X$ の直交基底である．さらに

$$\mathbf{u}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}, \quad j = 1, \cdots, s$$

とおけば，$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ は $X$ の正規直交基底になる．

シュミットの直交化より次のことがわかる．

**定理 10.9**　$X$ を $\boldsymbol{K}^N$ の非自明な部分空間とする．このとき $X$ の正規直交基底が存在する．

**証明．** $X$ の基底にシュミットの直交化を施せば，$X$ の正規直交基底が得られる．

■

**定理 10.10**　$X$ を $\boldsymbol{K}^N$ の部分空間で, $r = \dim X$ とする. $s < r$ とし, $\mathbf{u}_1, \cdots, \mathbf{u}_s \in X$ が $X$ の正規直交系とする. このとき, ある $\mathbf{u}_{s+1}, \cdots, \mathbf{u}_r \in X$ を選んで, $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ が $X$ の正規直交基底となるようにできる.

**証明.** 定理 9.19 より, $\mathbf{v}_{s+1}, \cdots, \mathbf{v}_r \in X$ を選んで, $\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{v}_{s+1}, \cdots, \mathbf{v}_r$ が $X$ の基底になるようにとれる. これにシュミットの直交化を適用すれば, $X$ の正規直交基底 $\{\mathbf{u}'_1, \cdots, \mathbf{u}'_r\}$ が得られる. シュミットの直交化の方法を見れば, $\mathbf{u}_1, \cdots, \mathbf{u}_s$ が正規直交系であると, $\mathbf{u}_i = \mathbf{u}'_i$, $i = 1, \cdots, s$ となっていることがわかる. よって求める正規直交基底が得られる. ■

## 10.4　直交射影と直交補空間

二つの集合 $X, Y \subset \boldsymbol{K}^N$ が直交しているとは, 任意の $\mathbf{x} \in X$ と任意の $\mathbf{y} \in Y$ に対して, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ となることである. このことを

$$X \perp Y$$

と表す.

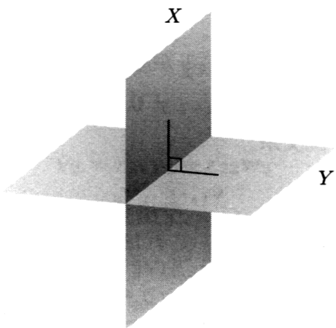



図 10.1

直交性も考慮に入れた部分空間の分解として, 次の直交分解の考え方がある.

**定義 10.6**　$S, X_1, X_2 \subset \boldsymbol{K}^N$ を $\boldsymbol{K}^N$ の部分空間で, $X_1, X_2 \subset S$ をみたしているとする. これらが次の二つの条件

(i)　任意の $\mathbf{x} \in S$ に対して, ある $\mathbf{x}_1 \in X_1$, $\mathbf{x}_2 \in X_2$ が存在し, $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$,

(ii)　$X_1 \perp X_2$

をみたすとき，$S$ は $X_1$ と $X_2$ に**直交分解**されるといい，

$$S = X_1 \oplus X_2$$

と表す．

直交分解は直和分解になっている．実際，$X_1 \cap X_2 \ni \mathbf{x}$ とすると，$\mathbf{x} \in X_1$, $\mathbf{x} \in X_2$ と $X_1 \perp X_2$ より，$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = 0$ が成り立っている．ゆえに $X_1 \cap X_2 = \{\mathbf{0}\}$ である．

**注意 10. 1**　本によっては，直和分解のほうを $S = X_1 \oplus X_2$ と表すものもある．本書では，直和分解は $S = X_1 \dot{+} X_2$ と表し，直交分解を $S = X_1 \oplus X_2$ と表す．

直交分解について次の定理は基本的なものである．

**定理 10. 11（直交分解定理）**　$S, X_1 \subset \boldsymbol{K}^N$ を $\boldsymbol{K}^N$ の部分空間とする．$X_1 \subset S$ であるとき，

$$S = X_1 \oplus X_2$$

をみたす部分空間 $X_2 \subset S$ が一意的に存在する．

**証明.** まず一意性を証明する．もし部分空間 $Y, Y' \subset S$ で $S = X_1 \oplus Y = X_1 \oplus Y'$ をみたすものが存在したとする．このとき，任意の $\mathbf{y} \in Y$ に対して，直交分解の定義から，$\mathbf{x} \in X_1$, $\mathbf{y}' \in Y'$ で

$$\mathbf{y} = \mathbf{x} + \mathbf{y}'$$

をみたすものが存在する．$X_1 \perp Y'$ より $\langle \mathbf{x}, \mathbf{y}' \rangle = 0$ が成り立っている．また，$X_1 \perp Y$ より $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ である．ゆえに

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} - \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y}' \rangle = 0.$$

したがって，$\mathbf{y} = \mathbf{y}' \in Y'$ となる．ゆえに $Y \subset Y'$ である．$Y$ と $Y'$ の役割を入れ替えて同様の議論をすれば，$Y' \subset Y$ も証明できる．したがって $Y = Y'$ が得られる．

次に $X_2$ の存在を証明する．$X_1 = \{\mathbf{0}\}$ の場合は $X_2 = S$，$X_1 = S$ の場合は $X_2 = \{\mathbf{0}\}$ とすればよい．$\dim X_1 = r < s = \dim S$ とする．定理 10.9 より，$X_1$ の正規直交基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ が存在する．定理 10.10 より，$\mathbf{u}_{r+1}, \cdots, \mathbf{u}_s \in S$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ が $S$ の正規直交基底となるようにとることができる．

$$X_2 = \operatorname{span}\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_s\}$$

とすれば，$X_1 \perp X_2$ である．さらに定理 10.8 より，任意の $\mathbf{x} \in S$ は，$\mathbf{x} = \sum_{j=1}^{s} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j$ と表せる．そこで

$$\mathbf{x}_1 = \sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j, \quad \mathbf{x}_2 = \sum_{j=r+1}^{s} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j$$

とすると，$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, $\mathbf{x}_1 \in X_1$ かつ $\mathbf{x}_2 \in X_2$ である．よって $S = X_1 \oplus X_2$ が成り立つ．■

**注意 10.2** 直交分解では $X_1$ を定めておけば分解の一意性が成り立つ．しかし直和分解については一意性は一般に成り立たない．たとえば $\boldsymbol{R}^2$ の部分空間として

$$X_1 = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} : x_1 \in \boldsymbol{R} \right\}, \quad X_2 = \left\{ \begin{pmatrix} 0 \\ x_2 \end{pmatrix} : x_2 \in \boldsymbol{R} \right\},$$

$$X_3 = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \boldsymbol{R} \right\}$$

を考える．このとき

$$\boldsymbol{R}^2 = X_1 \dotplus X_2 = X_1 \dotplus X_3$$

となっている．実際，明らかに $\boldsymbol{R}^2 = X_1 \dotplus X_2$ であり，また $X_1 \cap X_3 = \{\mathbf{0}\}$ および任意の $\mathbf{x} = (x_j)_{j=1}^2 \in \boldsymbol{R}^2$ に対して

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ x_2 \end{pmatrix}$$

であることから $\boldsymbol{R}^2 = X_1 \dotplus X_3$ でもある．なお $\boldsymbol{R}^2 = X_1 \dotplus X_2$ は直交分解でもあるが，$\boldsymbol{R}^2 = X_1 \dotplus X_3$ は直交分解にはなっていない．

**定義 10.7** 定理 10.11 における $X_2$ を $X_1$ の $S$ における**直交補空間**という．

$$X_2 = S \ominus X_1$$

と表す．特に $\boldsymbol{K}^N$ の部分空間 $X$ の $\boldsymbol{K}^N$ における直交補空間を $X$ の直交補空間といい，$X^{\perp}$ あるいは $(X)^{\perp}$ と表す．

定義から $X$ が $\boldsymbol{K}^N$ の部分空間のとき

$$\boldsymbol{K}^N = X \oplus X^{\perp}$$

である．定理 10.11 の証明から次のことがわかる．

**系 10.12**　$S, X_1, X_2$ を $\boldsymbol{K}^N$ の部分空間で，$X_1, X_2 \subset S$, $S = X_1 \oplus X_2$ とし，

$$\dim S = s, \quad \dim X_1 = r, \quad r < s$$

とする．$\mathbf{u}_1, \cdots, \mathbf{u}_r$ を $X_1$ の正規直交基底とし，$\{\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}. \cdots, \mathbf{u}_s\}$ を $S$ の正規直交基底であるとすると，$\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_s\}$ は $X_2$ の正規直交基底である．

**系 10.13**　$\left(\boldsymbol{K}^N\right)^\perp = \{\mathbf{0}\}$ である．また $X$ を $\boldsymbol{K}^N$ の部分空間とすると

$$\left(X^\perp\right)^\perp = X$$

である．

**証明.** $\boldsymbol{K}^N = \boldsymbol{K}^N \oplus \{\mathbf{0}\}$ より前半の主張は明らかである．後半の主張を示す．$\left(X^\perp\right)^\perp$ が $X^\perp$ の直交補空間であるから，$\boldsymbol{K}^N = X^\perp \oplus \left(X^\perp\right)^\perp$ である．一方 $\boldsymbol{K}^N = X \oplus X^\perp$ でもあるが，$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ より，$\boldsymbol{K}^N = X^\perp \oplus X$ でもある．したがって $\left(X^\perp\right)^\perp = X$ である．■

$S, X_1$ を $\boldsymbol{K}^N$ の部分空間で，$X_1 \subset S$ とする．このとき直交分解定理から，任意の $\mathbf{x} \in S$ に対して，ある $\mathbf{x}_1 \in X_1$ と $\mathbf{x}_2 \in S \ominus X_1$ で $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ となるものが一意的に存在する．そこで $\mathbf{x} \in S$ に対して $\mathbf{x}_1 \in X_1$ を対応させる写像を $P_{S,X_1}$ と表す．

**定義 10.8**　$P_{S,X_1}$ を $S$ から $X_1$ への**直交射影**という．特に $\boldsymbol{K}^N$ からその部分空間 $X$ への直交射影 $P_{\boldsymbol{K}^N,X}$ を，略して $P_X$ と表す．

**定理 10.14**　$S, X_1$ を $\boldsymbol{K}^N$ の非自明な部分空間で，$X_1 \subset S$ とする．$\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ を $X_1$ の正規直交基底とする．このとき，$\mathbf{x} \in S$ に対して，

$$P_{S,X_1}(\mathbf{x}) = \sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j$$

である．

**証明.** $s = \dim S$ とおく．$r < s$ の場合を示せば十分である．定理 10.10 より，$\mathbf{u}_{r+1}, \cdots, \mathbf{u}_s \in S$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ が $S$ の正規直交基底となるようにとることができる．このとき，定理 10.11 の証明中の記号を用いれば

$$\mathbf{x}_1 = \sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{u}_j \rangle \mathbf{u}_j$$

である．これより定理が得られる．■

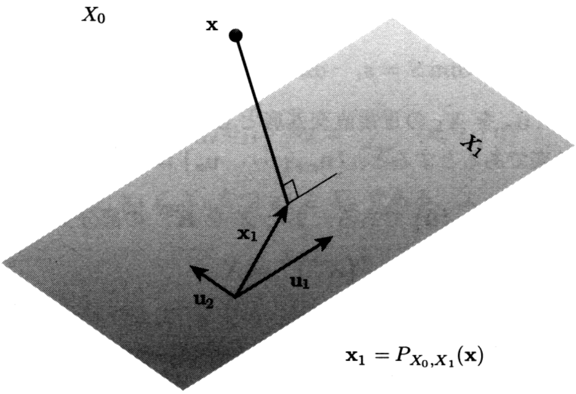



図 10.2

この定理の幾何的な解釈は，ベクトル $\mathbf{x} \in S$ を部分空間 $X_1$ に直交射影したものが，$X_1$ の正規直交基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ を一つの直交座標系としてみたとき，その第 $j$ 成分が内積 $\langle \mathbf{x}, \mathbf{u}_j \rangle$ によって与えられるということである (第 13 章参照).

**注意 10.3** $P_{S,X_1}$ の定義は正規直交基底のとり方によらずに定義されている．したがって，$\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ および $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ が $X_1$ の正規直交基底のとき

$$\sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{u}_j \rangle \, \mathbf{u}_j = P_{S,X_1}(\mathbf{x}) = \sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{v}_j \rangle \, \mathbf{v}_j$$

となっている．

**系 10.15** $S, X_1, X_2$ を $\boldsymbol{K}^N$ の部分空間で，$S = X_1 \oplus X_2$ であるとする．このとき，$Id_S - P_{S,X_1}$ は $S$ から $X_2$ への直交射影になっている．

**証明.** 定理 10.14 の証明中に定めた記号を用いると，任意の $\mathbf{x} \in S$ に対して

$$(Id_S - P_{S,X_1})(\mathbf{x}) = \sum_{j=r+1}^{s} \langle \mathbf{x}, \mathbf{u}_j \rangle \, \mathbf{u}_j$$

である．また系 10.12 より $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_s\}$ は $X_2$ の正規直交基底であるから，

$$P_{S,X_2}(\mathbf{x}) = \sum_{j=r+1}^{s} \langle \mathbf{x}, \mathbf{u}_j \rangle \, \mathbf{u}_j$$

であり，したがって $Id_S - P_{S,X_1} = P_{S,X_2}$ が得られる．■

直交補空間の次の特徴づけは有用である．

**系 10.16**　$S, X_1, X_2$ を $\boldsymbol{K}^N$ の部分空間で，$X_1, X_2 \subset S$, $S = X_1 \oplus X_2$ であるとする．このとき，

$$X_2 = \{\mathbf{x} : \mathbf{x} \in S \text{ かつ任意の } \mathbf{x}' \in X_1 \text{ に対して } \langle \mathbf{x}, \mathbf{x}' \rangle = 0\}$$

**証明.** 右辺を $Y$ とおく．$S = X_1 \oplus Y$ を示せば，直交分解定理より $Y = X_2$ である．任意に $\mathbf{x} \in S$ をとる．$S = X_1 \oplus X_2$ よりある $\mathbf{x}_1 \in X_1$ と $\mathbf{x}_2 \in X_2$ が存在し，$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ が成り立つ．$X_1 \perp X_2$ より任意の $\mathbf{x}' \in X_1$ に対して $\langle \mathbf{x}_2, \mathbf{x}' \rangle = 0$ をみたしているので，$\mathbf{x}_2 \in Y$ である．$Y$ の定義から明らかに $X_1 \perp Y$ も成り立っている．よって $S = X_1 \oplus Y$ である．■

定理 10.14 より $P_X$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像になっていることがわかる．したがってこれを表現する行列が存在する．次のような定義をしておく．

**定義 10.9**　$\boldsymbol{K}^N$ の部分空間 $X$ への直交射影 $P_X$ の表現行列を $X$ の上への**直交射影行列**という．また $A \in \boldsymbol{K}^{N \times N}$ が直交射影行列とは，$\boldsymbol{K}^N$ のある部分空間 $X$ の上への**直交射影行列**になっていることである．

次のことが成り立つ．

**定理 10.17**　$X$ を $\boldsymbol{K}^N$ の $r$ 次元部分空間とする $(r \geq 1)$．$A$ を $X$ の上への直交射影行列とする．このとき $\boldsymbol{K}^N$ の正規直交基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ をとり，$U = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$ が

$$U^* A U = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

をみたすようにできる ($U$ は後述の定義 13.6 のユニタリー行列である)．

**証明.** $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ を $X$ の正規直交基底で，$\mathbf{u}_{r+1}, \cdots, \mathbf{u}_N \in \boldsymbol{K}^N$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の正規直交基底となるようにとる．

$$U = (\mathbf{u}_1, \cdots, \mathbf{u}_N), \quad Q = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \in \boldsymbol{K}^{N \times N}$$

とする．このとき任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$U^* \mathbf{x} = \begin{pmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_N^* \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{u}_1^* \mathbf{x} \\ \vdots \\ \mathbf{u}_N^* \mathbf{x} \end{pmatrix}.$$

また，$UQ = (\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{0}, \cdots, \mathbf{0})$ である．したがって，

$$UQU^*\mathbf{x} = \sum_{j=1}^{r} (\mathbf{u}_j^*\mathbf{x})\, \mathbf{u}_j = \sum_{j=1}^{r} \langle \mathbf{x}, \mathbf{u}_j \rangle\, \mathbf{u}_j = P_X\mathbf{x} = A\mathbf{x}$$

となっている．ゆえに $UQU^* = A$ が成り立つ．また，$UU^* = U^*U = I_N$ であることが容易にわかる (後述の定理 13.16 参照)．よって，

$$U^*AU = U^*UQUU^* = Q$$

が得られる．■

$\boldsymbol{K}^N$ の部分空間 $X$ が与えられたとき，$X$ の上への直交射影行列はどのように求めることができるだろうか？　一つの作り方として次のようなものがある．

**系 10.18**　$X$ を $\boldsymbol{K}^N$ の $r$ 次元部分空間とし，$\mathbf{u}_1, \cdots, \mathbf{u}_r$ を $X$ の正規直交基底とする．

$$T = (\mathbf{u}_1, \cdots, \mathbf{u}_r) \quad (\in \boldsymbol{K}^{N\times r})$$

とすると，$TT^*$ は $\boldsymbol{K}^N$ から $X$ の上への直交射影行列である．また

$$TT^* = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_r\mathbf{u}_r^*$$

である．

**証明.** この証明では定理 10.17 の証明の記号を用いる．明らかに

$$UQ = \begin{pmatrix} T & O_{N\times(N-r)} \end{pmatrix},$$

$$QU^* = (UQ)^* = \begin{pmatrix} T^* \\ O_{(N-r)\times N} \end{pmatrix}$$

である．$QQ = Q$ より

$$A = UQU^* = UQQU^* = \begin{pmatrix} T & O_{N\times(N-s)} \end{pmatrix} \begin{pmatrix} T^* \\ O_{(N-s)\times N} \end{pmatrix}$$
$$= TT^*$$

が得られる．後半の主張は容易に示すことができる．■

直交射影行列は次のように特徴付けられる．

**定理 10.19**　$A \in \boldsymbol{K}^{N\times N}$ とする．このとき次の (1), (2) は同値である．

(1)　$A$ は直交射影行列である．

(2)　$AA = A$ かつ $A^* = A$ が成り立つ.

**証明.** (1) ⇒ (2)　定理 10.17 で定めた行列 $U$ により

$$U^*AU = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = Q$$

が成り立っている．したがって,

$$AA = UQU^*UQU^* = UQQU^* = UQU^* = A,$$
$$A^* = (UQU^*)^* = UQ^*U^* = UQU^* = A.$$

(2) ⇒ (1)　$\mathcal{R}(A)$ は $\boldsymbol{K}^N$ の部分空間である．任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，$\mathbf{x}_1 = A\mathbf{x}$ とおき，$\mathbf{x}_2 = (I - A)\,\mathbf{x}$ とおく．明らかに $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ かつ $\mathbf{x}_1 \in \mathcal{R}(A)$ である．以下では $\mathbf{x}_2$ が $\mathcal{R}(A)$ の任意の要素 $\mathbf{y}$ と直交していることを証明する．もしこのことが証明できれば，$\mathbf{x}_2 \in \mathcal{R}(A)^\perp$ であるから，定義より $A\mathbf{x} = P_{\mathcal{R}(A)}(\mathbf{x})$ となっていることがわかり，(1) が証明される．任意の $\mathbf{y} \in \mathcal{R}(A)$ をとる．定義より $\mathbf{y} = A\mathbf{x}'$ なる $\mathbf{x}' \in \boldsymbol{K}^N$ が存在する．したがって,

$$\begin{aligned}\langle \mathbf{y}, \mathbf{x}_2 \rangle &= \langle A\mathbf{x}', (I - A)\,\mathbf{x} \rangle = \langle A\mathbf{x}', \mathbf{x} \rangle - \langle A\mathbf{x}', A\mathbf{x} \rangle \\ &= \langle AA\mathbf{x}', \mathbf{x} \rangle - \langle A\mathbf{x}', A\mathbf{x} \rangle = \langle A\mathbf{x}', A^*\mathbf{x} \rangle - \langle A\mathbf{x}', A\mathbf{x} \rangle \\ &= \langle A\mathbf{x}', A\mathbf{x} \rangle - \langle A\mathbf{x}', A\mathbf{x} \rangle = 0\end{aligned}$$

が成り立っている．■

**定義 10.10**　$A \in \boldsymbol{K}^{N \times N}$ とする．$AA = A$ が成り立つとき，$A$ を**射影行列**という．

定理 10.19 は射影行列でかつエルミート行列であるものが直交射影行列であることを示している．

直交射影行列に対して次の不等式が成り立つ.

**定理 10.20**　$A \in \boldsymbol{K}^{N \times N}$ を直交射影行列とすると，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$\|A\mathbf{x}\| \leq \|\mathbf{x}\|.$$

**証明.** $\|A\mathbf{x}\| = 0$ の場合は明らかであるから，$\|A\mathbf{x}\| > 0$ の場合を証明する．このとき定理 10.19 より

$$\begin{aligned}\|A\mathbf{x}\|^2 &= \langle A\mathbf{x}, A\mathbf{x} \rangle = \langle A^*A\mathbf{x}, \mathbf{x} \rangle = \langle AA\mathbf{x}, \mathbf{x} \rangle \\ &= \langle A\mathbf{x}, \mathbf{x} \rangle\end{aligned}$$

である．したがって，シュヴァルツの不等式より

$$\|A\mathbf{x}\|^2 = \langle A\mathbf{x}, \mathbf{x}\rangle = |\langle A\mathbf{x}, \mathbf{x}\rangle| \leq \|A\mathbf{x}\| \, \|\mathbf{x}\|.$$

両辺を $\|A\mathbf{x}\|$ で割れば求める不等式が得られる． ■

**問題 10.3** $P = \dfrac{1}{N} 1_{N\times N}$ とする ($1_{N\times N}$ の定義は第 2.3 節の 5 参照)．また $Q = I_N - P$ とする．$P, Q$ は直交射影行列であることを示せ．さらに，

$$\mathcal{R}(Q) = \left\{ \mathbf{x} \in \boldsymbol{K}^N : \sum_{j=1}^{N} x_j = 0 \right\}$$

を示せ．

## 10.5 最良近似への応用

$X \subset \boldsymbol{K}^N$ とする．$\mathbf{u} \in \boldsymbol{K}^N$ に対して，$\mathbf{u}$ と $X$ との最短距離，すなわち

$$E_X(\mathbf{u}) = \inf \{ \|\mathbf{u} - \mathbf{x}\| : \mathbf{x} \in X \} \tag{10.6}$$

を $\mathbf{u}$ の $X$ に関する**近似誤差**という (inf は下限を表す記号である．定義は微分積分の教科書，たとえば [高木] を参照)．特に $X$ の元 $\mathbf{z}$ で

$$\|\mathbf{u} - \mathbf{z}\| = E_X(\mathbf{u})$$

をみたすものが存在するとき，$\mathbf{z}$ を $\mathbf{u}$ の $X$ に関する**最良近似**という．本書では $X$ が $\boldsymbol{K}^N$ の部分空間である場合の最良近似の求め方およびグラム行列を用いた近似誤差 $E_X(\mathbf{u})$ の計算方法を述べる．

**定理 10.21** $X$ を $\boldsymbol{K}^N$ の非自明な部分空間とする．このとき，任意の $\mathbf{u} \in \boldsymbol{K}^N$ に対して，$\mathbf{u}$ の $X$ に関する最良近似 $\mathbf{z}$ が一意的に存在し，それは $\mathbf{z} = P_X(\mathbf{u})$ で与えられる．すなわち $\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ を $X$ の正規直交基底とするとき

$$\mathbf{z} = \sum_{j=1}^{s} \langle \mathbf{u}, \mathbf{u}_j \rangle \, \mathbf{u}_j$$

で与えられる．

**証明.** 任意に $\mathbf{x} \in X$ をとる．$\mathbf{u} - P_X(\mathbf{u}) \in X^{\perp}$ かつ $P_X(\mathbf{u}) - \mathbf{x} \in X$ であるから，ピタゴラスの定理より

$$\begin{aligned}\|\mathbf{u}-\mathbf{x}\|^2 &= \|(\mathbf{u}-P_X(\mathbf{u}))+(P_X(\mathbf{u})-\mathbf{x})\|^2 \\ &= \|\mathbf{u}-P_X(\mathbf{u})\|^2+\|P_X(\mathbf{u})-\mathbf{x}\|^2 \\ &\geq \|\mathbf{u}-P_X(\mathbf{u})\|^2 .\end{aligned} \tag{10.7}$$

ここで $\|\mathbf{u}-P_X(\mathbf{u})\|^2$ は $\mathbf{x}$ のとり方に依存しない量であるから，

$$\inf\{\|\mathbf{u}-\mathbf{x}\| : \mathbf{x}\in X\} \geq \|\mathbf{u}-P_X(\mathbf{u})\|$$

が成り立つ．ゆえに

$$E_X(\mathbf{u}) \geq \|\mathbf{u}-P_X(\mathbf{u})\|$$

である．一方 $P_X(\mathbf{u})\in X$ であるから，明らかに

$$\|\mathbf{u}-P_X(\mathbf{u})\| \geq E_X(\mathbf{u})$$

も成り立っている．ゆえに $P_X(\mathbf{u})$ が $\mathbf{u}$ の $X$ に関する最良近似となっている．次に最良近似の一意性を証明しよう．もし $\mathbf{z}\in X$ が $\mathbf{u}$ の $X$ に関する最良近似であるとする．$P_X(\mathbf{u})$ も $\mathbf{u}$ の $X$ に関する最良近似であるから，(10.7) の計算を用いて

$$\begin{aligned}E_X(\mathbf{u})^2 = \|\mathbf{u}-\mathbf{z}\|^2 &= \|\mathbf{u}-P_X(\mathbf{u})\|^2+\|P_X(\mathbf{u})-\mathbf{z}\|^2 \\ &= E_X(\mathbf{u})^2+\|P_X(\mathbf{u})-\mathbf{z}\|^2\end{aligned}$$

を得る．したがって，$\|P_X(\mathbf{u})-\mathbf{z}\|=0$，すなわち $P_X(\mathbf{u})=\mathbf{z}$ となる．■

$X$ を $\boldsymbol{K}^N$ の $s$ 次元部分空間で，$\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ をその基底とするとき，この基底を用いて最良近似と近似誤差 $E_X(\mathbf{u})$ を求める方法を述べる．

**定理 10. 22**　$X$ を $\boldsymbol{K}^N$ の $s$ 次元部分空間 $(1\leq s<N)$ とし，$\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ をその基底とする．

(1)　$\mathbf{u}\in\boldsymbol{K}^N$ に対して

$$\begin{pmatrix} a_1 \\ \vdots \\ a_s \end{pmatrix} = G(\mathbf{u}_1,\cdots,\mathbf{u}_s)^{-1}\begin{pmatrix}\langle\mathbf{u},\mathbf{u}_1\rangle \\ \vdots \\ \langle\mathbf{u},\mathbf{u}_s\rangle\end{pmatrix} \tag{10.8}$$

(ただし $G(\mathbf{u}_1,\cdots,\mathbf{u}_s)$ はグラム行列) とおく．このとき

$$P_X(\mathbf{u}) = \sum_{j=1}^{s} a_j\mathbf{u}_j$$

が成り立つ．

(2) $\mathbf{u} \in \boldsymbol{K}^N$ に対して

$$E_X(\mathbf{u}) = \sqrt{\frac{\det G(\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{u})}{\det G(\mathbf{u}_1, \cdots, \mathbf{u}_s)}}.$$

**証明.** (1) $P_X(\mathbf{u}) \in \operatorname{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ より，$P_X(\mathbf{u}) = \sum_{k=1}^{s} a_k \mathbf{u}_k$ と表せるが，$\mathbf{u} - P_X(\mathbf{u}) \in X^\perp$ であるから，$j = 1, \cdots, s$ に対して，

$$\left\langle \mathbf{u} - \sum_{k=1}^{s} a_k \mathbf{u}_k, \mathbf{u}_j \right\rangle = 0.$$

ゆえに $\sum_{k=1}^{s} \langle \mathbf{u}_k, \mathbf{u}_j \rangle a_k = \langle \mathbf{u}, \mathbf{u}_j \rangle$ となっている．これより

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{u}_1, \mathbf{u}_s \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_s \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_s \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{u}_s \rangle \end{pmatrix} \tag{10.9}$$

が成り立つ．ここで

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{u}_1, \mathbf{u}_s \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_s \rangle \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^* \mathbf{u}_1 & \cdots & \mathbf{u}_1^* \mathbf{u}_s \\ \vdots & \ddots & \vdots \\ \mathbf{u}_s^* \mathbf{u}_1 & \cdots & \mathbf{u}_s^* \mathbf{u}_s \end{pmatrix} = G(\mathbf{u}_1, \cdots, \mathbf{u}_s)$$

であるが，系 9.10 より，$\det G(\mathbf{u}_1, \cdots, \mathbf{u}_s) > 0$ である．したがって (10.8) を得る．

(2) $\mathbf{u} - P_X(\mathbf{u}) \in X^\perp$ であるから，$\langle \mathbf{u} - P_X(\mathbf{u}), P_X(\mathbf{u}) \rangle = 0$ である．ゆえに

$$\begin{aligned} E_X(\mathbf{u})^2 &= \|\mathbf{u} - P_X(\mathbf{u})\|^2 = \langle \mathbf{u} - P_X(\mathbf{u}), \mathbf{u} - P_X(\mathbf{u}) \rangle \\ &= \langle \mathbf{u} - P_X(\mathbf{u}), \mathbf{u} \rangle = \left\langle \mathbf{u} - \sum_{k=1}^{s} a_k \mathbf{u}_k, \mathbf{u} \right\rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - \sum_{k=1}^{s} \langle \mathbf{u}_k, \mathbf{u} \rangle a_k. \end{aligned}$$

したがって (10.9) と合わせて，

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_1 \rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle \mathbf{u}_1, \mathbf{u}_s \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u}_s \rangle & 0 \\ \langle \mathbf{u}_1, \mathbf{u} \rangle & \cdots & \langle \mathbf{u}_s, \mathbf{u} \rangle & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ E_X(\mathbf{u})^2 \end{pmatrix} = \begin{pmatrix} \langle \mathbf{u}, \mathbf{u}_1 \rangle \\ \vdots \\ \langle \mathbf{u}, \mathbf{u}_s \rangle \\ \langle \mathbf{u}, \mathbf{u} \rangle \end{pmatrix}$$

を得る．$E_X(\mathbf{u})^2$ にクラメルの公式を適用すると

$$E_X(\mathbf{u})^2 = \frac{\det G(\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{u})}{\det G(\mathbf{u}_1, \cdots, \mathbf{u}_s)}$$

が得られる．■

この定理の系をいくつか述べておく．

**系 10.23**　$\mathbf{u}_1, \cdots, \mathbf{u}_n \in \boldsymbol{K}^N$ とする．このとき

$$\det G(\mathbf{u}_1, \cdots, \mathbf{u}_n) \leq \|\mathbf{u}_1\|^2 \cdots \|\mathbf{u}_n\|^2 .$$

**証明.** 系 9.10 より $\mathbf{u}_1, \cdots, \mathbf{u}_n$ が線形独立の場合を示せば十分である．$X_k = \operatorname{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ $(k = 1, \cdots, n)$ とする．定理 10.22 より

$$\begin{aligned}\det G(\mathbf{u}_1, \cdots, \mathbf{u}_n) &= E_{X_{n-1}}(\mathbf{u}_n)^2 \det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-1})\\ &\leq \|\mathbf{u}_n - \mathbf{0}\|^2 \det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-1})\\ &= \|\mathbf{u}_n\|^2 \det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-1}) .\end{aligned}$$

さらに

$$\begin{aligned}\det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-1}) &= E_{X_{n-2}}(\mathbf{u}_{n-1}) \det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-2})\\ &\leq \|\mathbf{u}_{n-1}\|^2 \det G(\mathbf{u}_1, \cdots, \mathbf{u}_{n-2})\end{aligned}$$

であるから，この議論を繰り返し行えば系を得る．最後の $G(\mathbf{u}_1)$ は定義から $\|\mathbf{u}_1\|^2$ である．■

この不等式から次の有名な不等式が得られる．

**系 10.24（アダマールの不等式）**　$A = (a_{jk}) \in \boldsymbol{K}^{N\times N}$ とする．このとき

$$|\det A|^2 \leq \prod_{k=1}^{N} \sum_{j=1}^{N} |a_{jk}|^2 .$$

**証明.** $\mathbf{a}_k = (a_{jk})_{j=1,\cdots,N} \in \boldsymbol{K}^N$ とおくと，系 7.7 と系 10.23 より

$$\begin{aligned}|\det A|^2 &= \det(A^*A) = \det G(\mathbf{a}_1, \cdots, \mathbf{a}_N)\\ &\leq \prod_{k=1}^{N} \|\mathbf{a}_k\|^2 = \prod_{k=1}^{N} \sum_{j=1}^{N} |a_{jk}|^2 .\end{aligned}$$

■

定理 10.22 を用いて次のことも証明できる．

**系 10.25** $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \boldsymbol{R}^3$ を線形独立なベクトルとする．この三つのベクトルから作られる平行六面体 $\Pi$ を考える (図 10.3 参照)．このとき，この平行六面体の体積を $V$ とすると

$$V = \sqrt{G(\mathbf{x}, \mathbf{y}, \mathbf{z})}.$$

**証明.** $L = \mathrm{span}\{\mathbf{x}\}$, $X = \mathrm{span}\{\mathbf{x}, \mathbf{y}\}$ とすると，$L$ は原点を通り，$\mathbf{x}$ に平行な直線である．また，$X$ は $\mathbf{x}, \mathbf{y}$ を含む平面を表している．このとき，$E_L(\mathbf{y})$ は $\mathbf{y}$ から直線 $\mathrm{span}\{\mathbf{x}\}$ への垂線の長さ，また $E_X(\mathbf{z})$ は $\mathbf{z}$ から平面 $\mathrm{span}\{\mathbf{x}, \mathbf{y}\}$ への垂線の長さであるから，

$$V = \|\mathbf{x}\| E_L(\mathbf{y}) E_X(\mathbf{z}) = \|\mathbf{x}\| \frac{\sqrt{G(\mathbf{x}, \mathbf{y})}}{\|\mathbf{x}\|} \frac{\sqrt{G(\mathbf{x}, \mathbf{y}, \mathbf{z})}}{\sqrt{G(\mathbf{x}, \mathbf{y})}} = \sqrt{G(\mathbf{x}, \mathbf{y}, \mathbf{z})}.$$

■

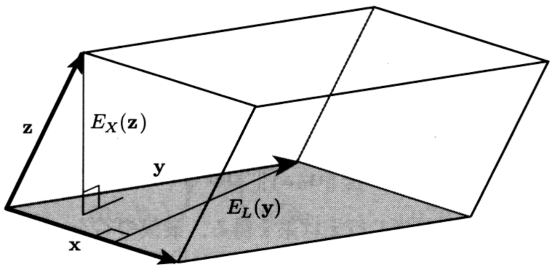



図 10.3

## 10.6 線形写像の値域と核

本節では線形写像の値域と核について，応用上も有用ないくつかの基本的な性質を証明する．

**定理 10.26** $A$ を $(M, N)$ 行列とする．このとき

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = N$$

が成り立つ．

**証明.** $\mathcal{N}(A) \subset \boldsymbol{K}^N$, $\mathcal{R}(A) \subset \boldsymbol{K}^M$ である．$s = \dim \mathcal{N}(A)$ とおく．$s \leq N$ である．$s = N$ ならば $\mathcal{N}(A) = \boldsymbol{K}^N$ であるから，$\mathcal{R}(A) = \{\mathbf{0}\}$，すなわち $\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = 0 + N = N$ が得られる．$1 \leq s < N$ の場合を示す．定

理 9.14 より $\mathcal{N}(A)$ の基底 $\mathbf{u}_1, \cdots, \mathbf{u}_s$ をとることができる．さらに定理 9.8 より $\mathbf{u}_{s+1}, \cdots, \mathbf{u}_N \in \boldsymbol{K}^N$ を $\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{u}_{s+1}, \cdots, \mathbf{u}_N$ が $\boldsymbol{K}^N$ の基底になるようにとることができる．$\{A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N\}$ が $\mathcal{R}(A)$ の基底になっていることを示す．任意の $A\mathbf{u} \in \mathcal{R}(A)$ に対して，$\mathbf{u} = \sum_{j=1}^{N} c_j \mathbf{u}_j$ が成り立つ．$\mathbf{u}_1, \cdots, \mathbf{u}_s \in \mathcal{N}(A)$ より

$$A\mathbf{u} = \sum_{j=1}^{N} c_j A\mathbf{u}_j = \sum_{j=s+1}^{N} c_j A\mathbf{u}_j \in \operatorname{span}\{A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N\}$$

である．

次に $A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N$ が線形独立であることを示す．$\sum_{j=s+1}^{N} \lambda_j A\mathbf{u}_j = \mathbf{0}$ とする．

$$A\left(\sum_{j=s+1}^{N} \lambda_j \mathbf{u}_j\right) = \sum_{j=s+1}^{N} \lambda_j A\mathbf{u}_j = \mathbf{0}$$

より $\sum_{j=s+1}^{N} \lambda_j \mathbf{u}_j \in \mathcal{N}(A)$ となっている．したがって，

$$\sum_{j=s+1}^{N} \lambda_j \mathbf{u}_j = \sum_{j=1}^{s} \mu_j \mathbf{u}_j$$

と表せる．すなわち

$$\sum_{j=1}^{s} (-\mu_j)\,\mathbf{u}_j + \sum_{j=s+1}^{N} \lambda_j \mathbf{u}_j = \mathbf{0}$$

である．ここで $\mathbf{u}_1, \cdots, \mathbf{u}_N$ は線形独立であるから，

$$\mu_1 = \cdots = \mu_s = \lambda_{s+1} = \cdots = \lambda_N = 0$$

が導かれ，$A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N$ が線形独立であることが得られる．

ゆえに $\{A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N\}$ が $\mathcal{R}(A)$ の基底である．よって

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = N - s + s = N$$

が成り立つ．

$s = 0$ の場合も同様の議論で示せる．■

**系 10.27（交代定理）** $T$ を $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像とする．次の (1), (2) は同値である．

(1)　$T$ は全射である．

(2)　$T$ は単射である．

**証明.** $T$ を表現する $(N,N)$ 行列を $A$ とすると，定理 10.26 より

$$\dim\mathcal{R}(T)+\dim\mathcal{N}(T)=\dim\mathcal{R}(A)+\dim\mathcal{N}(A)=N$$

である．もし $T$ が全射ならば $\dim\mathcal{R}(T)=\dim\boldsymbol{K}^N=N$ より $\dim\mathcal{N}(T)=0$ であるから，命題 3.3 より $T$ は単射であることがわかる．一方，$T$ が単射ならば $\dim\mathcal{N}(T)=0$ より $\dim\mathcal{R}(T)=N$ であるから，命題 9.16 より $\mathcal{R}(T)=\boldsymbol{K}^N$ となり，$T$ が全射となる． ■

$A$ と $A^*$ の値域と核には次の関係がある．

**定理 10.28**　(1)　$\mathcal{N}(A)=\mathcal{R}(A^*)^\perp$,　$\mathcal{N}(A^*)=\mathcal{R}(A)^\perp$.

(2)　$\mathcal{R}(A)=\mathcal{N}(A^*)^\perp$,　$\mathcal{R}(A^*)=\mathcal{N}(A)^\perp$.

**証明.** (1)　$\mathbf{x}\in\mathcal{N}(A)$ とする．任意の $\mathbf{y}\in\mathcal{R}(A^*)$ を考える．このときある $\mathbf{z}\in\boldsymbol{K}^M$ で $\mathbf{y}=A^*\mathbf{z}$ となるものが存在する．したがって，

$$\langle\mathbf{x},\mathbf{y}\rangle=\langle\mathbf{x},A^*\mathbf{z}\rangle=\langle A\mathbf{x},\mathbf{z}\rangle=\langle\mathbf{0},\mathbf{z}\rangle=0.$$

ゆえに $\mathbf{x}\in\mathcal{R}(A^*)^\perp$ である．逆に $\mathbf{x}\in\mathcal{R}(A^*)^\perp$ とする．任意の $\mathbf{y}\in\boldsymbol{K}^M$ に対して

$$\langle A\mathbf{x},\mathbf{y}\rangle=\langle\mathbf{x},A^*\mathbf{y}\rangle=0$$

が成り立つ．したがって，命題 10.1 (2) より $A\mathbf{x}=\mathbf{0}$，すなわち $\mathbf{x}\in\mathcal{N}(A)$ となる．

後半の主張は $A^{**}=A$ より明らか．

(2)　$\mathcal{R}(A)=\left(\mathcal{R}(A)^\perp\right)^\perp=\mathcal{N}(A^*)^\perp$ である．残りの主張は $A^{**}=A$ より明らか． ■

**系 10.29**　$A\in\boldsymbol{K}^{M\times N}$ とする．このとき

$$\boldsymbol{K}^N=\mathcal{R}(A^*)\oplus\mathcal{N}(A).$$

**証明.** 定理 10.28 より，

$$\boldsymbol{K}^N=\mathcal{R}(A^*)\oplus\mathcal{R}(A^*)^\perp=\mathcal{R}(A^*)\oplus\mathcal{N}(A).$$

■

**問題 10.4**　任意の $N$ 次正方行列 $A$ に対して，

$$\boldsymbol{K}^N = \mathcal{R}(A) \dotplus \mathcal{N}(A)$$

は成り立っているか？

**問題 10.5**　問題 10.3 の $Q$ について，$\dim R(Q)$ を求めよ．

**問題 10.6**　$A \in \boldsymbol{K}^{N \times M}$ から定まる線形写像 $T_A$ に対して，$T_A$ の $\mathcal{R}(A^*)$ への制限 $T_A|_{\mathcal{R}(A^*)}$ は $\mathcal{R}(A^*)$ から $\mathcal{R}(A)$ への全単射である．したがって逆写像 $(T_A|_{\mathcal{R}(A^*)})^{-1}$ が存在し，$(T_A|_{\mathcal{R}(A^*)})^{-1}$ は $\mathcal{R}(A)$ から $\mathcal{R}(A^*)$ への全単射かつ線形写像である．

**問題 10.7**　$A \in \boldsymbol{K}^{N \times M}$, $B \in \boldsymbol{K}^{N \times L}$ とし，$\mathcal{R}(A) = \mathcal{R}(B)$ であるとする．このときある $C \in \boldsymbol{K}^{M \times L}$ で

$$AC = B$$

をみたすものが存在する．

# 第11章　行列の階数

$(M, N)$ 行列 $A$ に対して，$\boldsymbol{K}^M$ の部分空間 $\mathcal{R}(A)$ の次元を $A$ の**階数**あるいは**ランク**といい，$\mathrm{rank}(A)$ と表す．すなわち

$$\mathrm{rank}(A) = \dim \mathcal{R}(A)$$

である．本章では行列の階数を求める方法を解説する．

**ポイント 11**　行列の階数の概念は連立 1 次方程式を深く考察する場合に必要不可欠になる．階数の定義および具体的な行列の階数を計算できるようにしておこう．

## 11.1　一般論

まず一般的な定理をいくつか証明する．

**定理 11.1**　$(M, N)$ 行列 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ に対して，

$$\mathcal{R}(A) = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$$

であり，

$$\mathrm{rank}(A) = \dim \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$$

である．

**証明.** $\mathbf{x} = (x_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ に対して，$A\mathbf{x} = \sum_{j=1}^{N} x_j \mathbf{a}_j \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ であるから，$\mathcal{R}(A) \subset \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ である．

逆に，$\sum_{j=1}^{N} x_j \mathbf{a}_j \in \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ に対して，$\mathbf{x} = (x_j)_{j=1,\cdots,N}$ とおけば，$\sum_{j=1}^{N} x_j \mathbf{a}_j = A\mathbf{x} \in \mathcal{R}(A)$ であるから，$\mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} \subset \mathcal{R}(A)$ である．■

**系 11.2**　$r < N$ とする．$(M, N)$ 行列 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ に対して，次の (1), (2) は同値である．

(1)　$r = \mathrm{rank}(A)$.

(2)　$\mathbf{a}_1, \cdots, \mathbf{a}_N$ の中のある $r$ 個のベクトルで線形独立なものが存在し，$r+1$ 個以上のベクトルはいかなるものも線形独立にはならない．

**証明.** (1) ⇒ (2)　定理 11.1 より $r = \dim \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ である．したがって，定理 9.17 より，$\mathbf{a}_1, \cdots, \mathbf{a}_N$ の中から $\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}$ を選んで $\mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ の基底になるようにとれる．このことから (2) が得られる．

(2) ⇒ (1)　$\dim \mathcal{R}(A) = \dim \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} = r$ である．■

**系 11.3**　$A$ を $(M, N)$ 行列とすると，$\mathrm{rank}(A) \leq \min\{M, N\}$.

**証明.**　$\mathcal{R}(A)$ は $\boldsymbol{K}^M$ の部分空間であるから，$\mathrm{rank}(A) \leq M$ である．また $\dim \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} \leq N$ であるから系が得られる．■

交代定理 (系 10.27) から次のことがわかる．

**定理 11.4**　$A$ を $N$ 次正方行列とする．このとき次の (1), (2), (3) は同値である．

(1)　$A$ は正則行列である．

(2)　$|A| \neq 0$.

(3)　$\mathrm{rank}(A) = N$.

**証明.** (1) ⇔ (2) はすでに証明した．$A$ で定義される線形写像を $T$ とする．(3) が成立すれば $\dim \mathcal{R}(A) = N$ より $\mathcal{R}(A) = \boldsymbol{K}^N$ となり，$T$ は全射である．ゆえに，交代定理より $T$ は単射でもある．したがって，$T$ の逆写像，すなわち $A$ の逆行列が存在する．すなわち (1) が成り立つ．逆に (1) が成り立てば，$T$ は全単射であるから $\mathrm{rank}(A) = \dim \mathcal{R}(T) = N$ が得られる．■

定理 9.5 より次のことが導かれる．

**定理 11.5** $\mathbf{a}_1, \cdots, \mathbf{a}_N \in \boldsymbol{K}^N$ とし $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．このとき次の (1), (2) は同値である．

(1) $\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ は $\boldsymbol{K}^N$ の基底である．

(2) $A$ は正則である．

定理 11.4 の一般化として，小行列式を用いて階数の特徴づけができることを示す．

**定理 11.6** $A$ を $(M, N)$ 行列とする．

(1) $1 \leq r < \min\{M, N\}$ とする．$\mathrm{rank}(A) = r$ であるための必要十分条件は，$A$ の $r$ 次の小行列式のなかに 0 でないものがあり，$r+1$ 次以上の小行列式がすべて 0 になることである．

(2) $s = \min\{M, N\}$ とする．$\mathrm{rank}(A) = s$ であるための必要十分条件は，$A$ の $s$ 次の小行列式の中に 0 でないものがあることである．

**証明.** (1) $r = \mathrm{rank}(A)$ とする．$X = \mathrm{span}\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ とおく．このとき $\dim X = r < N$ であるから，この中から $r$ 個のベクトル $\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}$ を選んで，$X$ の基底にできる．したがって，定理 9.9 より $(\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)})$ の $r$ 次の小行列式で 0 でないものが存在する．さて，もしも $A$ の $r+1$ 次以上の小行列式で 0 でないものが存在したとする．その小行列式を

$$\begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_n} \\ \vdots & \ddots & \vdots \\ a_{i_n j_1} & \cdots & a_{i_n j_n} \end{vmatrix}$$

$(n \geq r+1)$ とする．このとき，定理 9.9 より $\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_n}$ は線形独立でなければならない．したがって，$\mathrm{rank}(A) = \dim X \geq n \geq r+1$ となり矛盾である．したがって，$r+1$ 次以上の小行列式がすべて 0 になる．

逆を示す．

$$\begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_r} \\ \vdots & \ddots & \vdots \\ a_{i_r j_1} & \cdots & a_{i_r j_r} \end{vmatrix} \neq 0$$

とする．このとき，$\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}$ は線形独立である．これ以外のベクトル $\mathbf{a}_k$ を任意にとる．このとき，仮定から $(\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}, \mathbf{a}_k)$ の任意の $r+1$ 次小行列式は 0 なので，$\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}, \mathbf{a}_k$ は線形従属である．ゆえに，$\mathbf{a}_k \in \mathrm{span}\{\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}\}$

であるから $X = \mathrm{span}\{\mathbf{a}_{j_1}, \cdots, \mathbf{a}_{j_r}\}$ となる．ゆえに $\mathrm{rank}(A) = \dim X = r$ である．

(2) は定理 9.9 と (1) より明らかである．■

**系 11.7**　$A$ を $(M, N)$ 行列とする．このとき，

$$\mathrm{rank}(A) = \mathrm{rank}(A^T) = \mathrm{rank}(A^*).$$

**証明.** $r$ 次正方行列 $B$ に対して，$\det B^T = \det B$ であることに注意する．$A$ の $r$ 次の小行列 $B$ に対して，$B^T$ は $A^T$ の $r$ 次の小行列である．逆も正しい．ゆえに，定理 11.6 より $A$ と $A^T$ の階数は一致する．$A^*$ は $A^T$ の各成分の複素共役をとったものであり，$r$ 次正方行列 $B$ に対して，$B$ の各成分の複素共役をとったものを $\overline{B}$ とすると $\det \overline{B} = \overline{\det B}$ であるから，定理 11.6 より，$A^T$ と $A^*$ の階数は一致する．■

**系 11.8**　$A$ を $(M, N)$ 行列とする．$r < \min\{M, N\}$ とする．次の (1), (2) は同値である．

(1)　$r = \mathrm{rank}(A)$.

(2)　$A(1, :)^T, \cdots, A(M, :)^T$ の中のある $r$ 個のベクトルで線形独立なものが存在し，かついかなる $r+1$ 個以上のベクトルも線形独立でない．

**証明.** $\mathrm{rank}(A^T) = \mathrm{rank}(A) = r$ である．$A^T(:, i) = A(i, :)^T$ であるから，$A^T$ に対して系 11.2 を用いればよい．■

**定理 11.9**　$A$ を $(L, M)$ 行列，$B$ を $(M, N)$ 行列とする．このとき，

$$\mathrm{rank}(A) + \mathrm{rank}(B) - M \leq \mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}.$$

**証明.** まず二番目の不等式を証明する．$\mathcal{R}(AB) \subset \mathcal{R}(A)$ より $\mathrm{rank}(AB) \leq \mathrm{rank}(A)$ が得られる．$\dim \mathcal{R}(B) = s$ とする．$\mathcal{R}(B)$ の基底を $\mathbf{u}_1, \cdots, \mathbf{u}_s \in \boldsymbol{K}^M$ とする．任意の $\mathbf{y} \in \mathcal{R}(AB)$ に対して，$\mathbf{y} = AB\mathbf{x}$ となる $\mathbf{x} \in \boldsymbol{K}^N$ が存在する．$B\mathbf{x} \in \mathcal{R}(B)$ より $B\mathbf{x} = \sum_{j=1}^{s} \alpha_j \mathbf{u}_j$ が成立している．したがって，$\mathbf{y} = A\left(\sum_{j=1}^{s} \alpha_j \mathbf{u}_j\right) = \sum_{j=1}^{s} \alpha_j A\mathbf{u}_j$．ゆえに，$\mathcal{R}(AB) \subset \mathrm{span}\{A\mathbf{u}_1, \cdots, A\mathbf{u}_s\}$．よって，$\mathrm{rank}(AB) \leq s = \mathrm{rank}(B)$ も成り立つ．

次に一番目の不等式を証明する．

$$\begin{aligned}
\operatorname{rank}(A)+\operatorname{rank}(B) &= \operatorname{rank}\begin{pmatrix} O_{L,N} & A \\ B & O_{M,M} \end{pmatrix} \leq \operatorname{rank}\begin{pmatrix} O_{L,N} & A \\ B & I_M \end{pmatrix} \\
&= \operatorname{rank}\begin{pmatrix} I_L & A \\ O_{M,L} & I_M \end{pmatrix}\begin{pmatrix} -AB & O_{L,M} \\ B & I_M \end{pmatrix} \\
&= \operatorname{rank}\begin{pmatrix} I_L & A \\ O_{M,L} & I_M \end{pmatrix}\begin{pmatrix} -AB & O_{L,M} \\ O_{M,N} & I_M \end{pmatrix}\begin{pmatrix} I_N & O_{N,M} \\ B & I_M \end{pmatrix} \\
&\leq \operatorname{rank}\begin{pmatrix} -AB & O_{L,M} \\ O_{M,N} & I_M \end{pmatrix} \\
&= \operatorname{rank}(AB)+M.
\end{aligned}$$

■

**定理 11.10** (1)　$A$ を $(M,N)$ 行列とし，$P$ を $M$ 次正則行列とする．このとき

$$\operatorname{rank}(PA)=\operatorname{rank}(A).$$

(2)　$A$ を $(M,N)$ 行列とし，$Q$ を $N$ 次正則行列とする．このとき

$$\operatorname{rank}(AQ)=\operatorname{rank}(A).$$

(3)　$A$ を $(M,N)$ 行列とし，$P$ を $M$ 次正則行列，$Q$ を $N$ 次正則行列とする．このとき，

$$\operatorname{rank}(PAQ)=\operatorname{rank}(A)$$

である．

**証明.** (1)　$\dim\mathcal{R}(A)=s$ とする．$\mathcal{R}(A)$ の基底を $\mathbf{u}_1,\cdots,\mathbf{u}_s\in\boldsymbol{K}^M$ とする．定理 11.9 の証明と同様にして $\mathcal{R}(PA)\subset\operatorname{span}\{P\mathbf{u}_1,\cdots,P\mathbf{u}_s\}$ であるが，この包含関係と逆向きの包含関係も明らかに成立するので，$\mathcal{R}(PA)=\operatorname{span}\{P\mathbf{u}_1,\cdots,P\mathbf{u}_s\}$ である．ここで，$P\mathbf{u}_1,\cdots,P\mathbf{u}_s$ は命題 9.2 (3) より線形独立である．よって

$$\operatorname{rank}(PA)=\dim\mathcal{R}(PA)=s=\operatorname{rank}(A).$$

(2) は問題とする (問題 11.1)．(3) は (1), (2) より明らかである．■

**問題 11.1**　定理 11.10 (2) を証明せよ．

**系 11. 11**　$A$ を $(M, N)$ 行列とする．$A$ を有限回の行に関する基本変形および列に関する基本変形により $A'$ に変形したとき，$\mathrm{rank}(A) = \mathrm{rank}(A')$ である．

**証明.** 仮定より，ある $M$ 次基本行列 $E_1, \cdots, E_k$ と $N$ 次基本行列 $F_1, \cdots, F_l$ が存在し，

$$E_k \cdots E_1 A F_1 \cdots F_l = A'$$

である．$E_k \cdots E_1$ も $F_1 \cdots F_l$ も正則行列なので定理 11.10 より系が得られる．■

**定理 11. 12**　$A$ を $(M, N)$ 行列とする．このとき

$$\mathrm{rank}(A) = \mathrm{rank}(AA^*) = \mathrm{rank}(A^*A).$$

**証明.** まず，$\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$A^*\mathbf{x} = \mathbf{0} \Longleftrightarrow AA^*\mathbf{x} = \mathbf{0} \tag{11.1}$$

を示す．もしこれが証明できれば，$\mathcal{N}(A^*) = \mathcal{N}(AA^*)$ であるから，定理 10.28 より，

$$\mathcal{R}(A) = \mathcal{N}(A^*)^{\perp} = \mathcal{N}(AA^*)^{\perp} = \mathcal{R}(AA^*)$$

である．ゆえに $\mathrm{rank}(A) = \mathrm{rank}(AA^*)$.

(11.1) の $\Rightarrow$ は明らかだから，$\Leftarrow$ を示す．$AA^*\mathbf{x} = \mathbf{0}$ とすると

$$\|A^*\mathbf{x}\|^2 = \langle A^*\mathbf{x}, A^*\mathbf{x}\rangle = \langle \mathbf{x}, AA^*\mathbf{x}\rangle = \langle \mathbf{x}, \mathbf{0}\rangle = 0$$

であるから，$A^*\mathbf{x} = \mathbf{0}$ である．

残りの主張を示すため，上記の議論を $A$ の代わりに $A^*$ を用いて行う．すると $A^{**} = A$ であるから $\mathrm{rank}(A^*) = \mathrm{rank}(A^*A)$ が得られる．また系 11.7 より $\mathrm{rank}(A) = \mathrm{rank}(A^*)$ である．■

**系 11. 13**　$A$ を $(M, N)$ 行列とする．$\mathrm{rank}(A) = M$ ならば $AA^*$ は正則であり，$\mathrm{rank}(A) = N$ ならば $A^*A$ は正則である．

**証明.** $\mathrm{rank}(A) = M$ の場合，定理 11.12 より $\mathrm{rank}(AA^*) = \mathrm{rank}(A) = M$ であり，$AA^*$ は $M$ 次正方行列である．したがって，定理 11.4 より $AA^*$ は正則である．残りの主張も同様にして証明できる．■

## 11.2 階数の計算とピボット

系 11.11 より行列の階数は有限回の行および列に関する基本変形を施しても変わらない．したがって行列が与えられたとき，その階数を計算する方法の一つとして，行列を基本変形により階数の計算しやすい簡単な形の行列に変形するということが考えられる．

はじめに階数の計算しやすい簡単な形の行列として，行階段行列と行既約階段行列について述べ，その階数の算出方法を解説する．次節において，すべての行列が，行に関する基本変形により行階段行列あるいは行既約階段行列に変形できることを示す (定理 11.19 参照)．

行列

$$\Sigma = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}$$

の第 $j$ 行 $a_{j1}, \cdots, a_{jN}$ を考える．この中に 0 でない成分が少なくとも一つ存在するとき，左から数えて最初の 0 でない成分を $\Sigma$ の第 $j$ 行の**ピボット** (pivot) という．すなわち，$k = \min\{l : a_{jl} \neq 0\}$ としたとき，$a_{jk}$ がピボットである．$a_{jk}$ がピボットであることを $\boxed{a_{jk}}$ で表す．もし $a_{j1}, \cdots, a_{jN}$ がすべて 0 である場合は，この行にピボットは存在しないという．

**例 11.1** $\Sigma = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ とする．第 1 行のピボットは 1，第 2 行のピボットは 3，第 3 行のピボットは存在しない．すなわち，

$$\begin{pmatrix} \boxed{1} & 0 & 2 & 0 \\ 0 & \boxed{3} & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

と表せる．

**定義 11.1** $(M, N)$ 行列 $\Sigma = (a_{jk})$ が**行階段行列** (row echelon form) であるとは，

(1) ピボットの存在しない行があれば，それはピボットの存在するどの行よりも下にある．

(2)　第 $j+1$ 行にピボットが存在するとき，第 $j$ 行のピボットは第 $j+1$ 行のピボットより左側の列にある．すなわち，第 $j+1$ 行のピボットが $a_{j+1,k}$ で，第 $j$ 行のピボットが $a_{j,k'}$ とすると，$k'<k$ である (このことは，$k>1$ であることも意味している).

特に行階段行列のうち，すべてのピボットが 1 であり，さらにピボットを含む列のピボット以外の成分がすべて 0 であるとき，**行既約階段行列** (reduced row echelon form) という．

たとえば，次の形の行列は行階段行列である．

$$R=\begin{pmatrix} \boxed{\triangle} & * & * & * & * & * & * & * & * \\ 0 & 0 & \boxed{\triangle} & * & * & * & * & * & * \\ 0 & 0 & 0 & \boxed{\triangle} & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \boxed{\triangle} & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\triangle} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\triangle} & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ここで $\triangle$ は 0 でない数，$*$ は適当な数を表す．また

$$R=\begin{pmatrix} \boxed{1} & * & 0 & 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & \boxed{1} & 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & \boxed{1} & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & \boxed{1} & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

は行既約階段行列である．一般に行階段行列は上三角行列である．

**例 11.2** $\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$ は行階段行列ではない．第 1 行のピボットが第 2 行のピボットより右側の列にある．しかし，1 行と 2 行を入れ替えた行列

$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$ は行階段行列である．しかし，行既約階段行列ではない．なぜなら第3行のピボットを含む列，すなわち第3列はピボット以外に0でない成分がある．$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$ は行既約階段行列である．

**例 11.3** $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ は行階段行列でない．定義 11.1 (1) が成り立っていない．しかし，2行と3行を入れ替えた行列 $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ は行既約階段行列である．

**例 11.4** $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ は行階段行列でない．第2行のピボットが第3行のピボットのある列より左側にない．

**定理 11.14** $A$ が行階段行列のとき，

$$\mathrm{rank}(A) = \text{“}A\text{ のピボットの数”}$$

**証明.** $A$ を $(M, N)$ 行階段行列とする．$A = (a_{ij})$ とし，ピボットを $a_{k,j(k)}$, $k = 1, \cdots, r$, $j(1) < \cdots < j(r)$ とする．$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ おく．$\mathbf{a}_{j(1)}, \cdots, \mathbf{a}_{j(r)}$ がピボットを含むベクトルである．第 $j(k)$ 列の $-\dfrac{a_{ki}}{a_{k,j(k)}}$ 倍を第 $i$ 列 $(i \neq j(k))$ に加えることにより，ピボットを含む行のピボット以外の成分をすべて0にすることができる (参考例参照)．さらに各 $k = 1, \cdots, r$ についてピボット $a_{k,j(k)}$ を含む行を $\dfrac{1}{a_{k,j(k)}}$ 倍することにより，すべてのピボットを1にできる．そしてさらに列の入れ替えにより $j(k)$ が $k$ になるような置換 $\sigma$ に対する置換行列 $P$(第 7.2

節参照) を用いて, $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ とできる.

$$\begin{pmatrix} \boxed{\triangle} & * & * & * & * & * \\ 0 & 0 & \boxed{\triangle} & * & * & * \\ 0 & 0 & 0 & \boxed{\triangle} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{\triangle} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{\triangle} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{\triangle} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

参考例

変形した後の行列を $A'$ とすると, $A' = (\mathbf{e}_1, \cdots, \mathbf{e}_r, \mathbf{0}, \cdots, \mathbf{0})$ であるから, 定理 11.1 より $\mathrm{rank}(A') = r$ である. また $A$ から $A'$ への変形は, 行に関する基本変形と列に関する基本変形によるものであるから, 系 11.11 より $\mathrm{rank}(A) = \mathrm{rank}(A') = r$ である. $r$ はピボットの数である. ■

**命題 11. 15**　行階段行列は行に関する基本変形により行既約階段行列にすることができる.

**証明.** 行階段行列 $A = (a_{ij}) \in \boldsymbol{K}^{M \times N}$ のピボットを $a_{k,j(k)}$, $k = 1, \cdots, r$, $j(1) < \cdots < j(r)$ とする. 第 $i$ 行 $(i = 1, \cdots, k-1)$ に第 $k$ 行の $-\dfrac{a_{i,j(k)}}{a_{k,j(k)}}$ 倍を加えることにより, $(i, j(k))$ 成分を 0 にすることができる. このようにするとピボットを含む列のピボット以外の成分はすべて 0 になる. 各 $k$ 行 $(k = 1, \cdots, r)$ を $\dfrac{1}{a_{k,j(k)}}$ 倍すれば, すべてのピボットが 1 になる. これらの行に関する基本変形により, 行階段行列は行既約階段行列になる. ■

次節において, 任意の行列が行に関する基本変形により行階段行列に変形する方法を示すが, それを用いれば, 行列の階数はその変形した行階段行列のピボットの数を数えればよいことがわかる.

## 11.3 行列の標準化

ここでは任意の $(M, N)$ 行列が行による基本変形で，行既約階段行列に変形できること，さらに列の基本変形を施すことにより，$r < M, N$ の場合は

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix},$$

また $r = M < N$ の場合，$r = N < M$ の場合はそれぞれ

$$\begin{pmatrix} I_M & O \end{pmatrix}, \quad \begin{pmatrix} I_N \\ O \end{pmatrix},$$

そして，$r = M = N$ の場合は $I_M$ のいずれかの形に変形できることを示す．なお，今後記述を簡単にするため，次の表記上の規約を設ける．

**規約 11.16**　$(M, N)$ 行列

$$\begin{pmatrix} I_r & K_1 \\ K_2 & K_3 \end{pmatrix}$$

により，$r = M < N$ の場合は $\begin{pmatrix} I_M & K_1 \end{pmatrix}$, $r = N < M$ の場合は $\begin{pmatrix} I_N \\ K_2 \end{pmatrix}$, $r = M = N$ の場合は $I_M$ を表すものとする．

**定理 11.17**　$A$ を $(M, N)$ 行列で，$r = \mathrm{rank}(A)$ とする．このとき有限個の $M$ 次基本行列 $E_1, \cdots, E_k$ と $N$ 次置換行列 $P$ により

$$E_k \cdots E_1 A P = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix}$$

(ここで $K$ はある $(r, N - r)$ 行列) と変形できる．

**証明.** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ と表す．$\mathbf{a}_1, \cdots, \mathbf{a}_N$ の中で，$\mathbf{0}$ でない最初のベクトルを $\mathbf{a}_{i_1}$ とする．このとき

$$A = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{a}_{i_1}, \mathbf{a}_{i_1+1}, \cdots, \mathbf{a}_N).$$

行に関する基本変形をして，$\mathbf{a}_{i_1}$ が $\mathbf{e}_1$ になるようにすることができる．その結果を

$$A' = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{e}_1, \mathbf{a}'_{i_1+1}, \cdots, \mathbf{a}'_N)$$

とおく．次に $\mathbf{a}'_{i_1+1}, \cdots, \mathbf{a}'_N$ の中で，$\mathbf{e}_1$ の定数倍で表せない最初のベクトルを $\mathbf{a}'_{i_2}$ とすると

$$A' = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{e}_1, \lambda_1 \mathbf{e}_1, \cdots, \lambda_{k_1} \mathbf{e}_1, \mathbf{a}'_{i_2}, \cdots, \mathbf{a}'_N)$$

となっている ($\lambda_i = 0$ ということもありうる)．ここで第 2 行と第 $j$ 行 ($j \geq 3$) を入れ替える，第 2 行の定数倍を他の行に加える，第 2 行を定数倍する，といった行に関する基本変形をして

$$A'' = (\mathbf{0}, \cdots, \mathbf{0}, \mathbf{e}_1, \lambda_1 \mathbf{e}_1, \cdots, \lambda_{k_1} \mathbf{e}_1, \mathbf{e}_2, \mathbf{a}''_{i_2+1}, \cdots, \mathbf{a}''_N)$$

とすることができる．$\mathbf{a}''_{i_2+1}, \cdots, \mathbf{a}''_N$ の中で $\mathbf{e}_1$ と $\mathbf{e}_2$ の線形結合で表せない最初のベクトルを $\mathbf{a}''_{i_3}$ とする．このとき

$$\begin{aligned} A'' = (&\mathbf{0}, \cdots, \mathbf{0}, \mathbf{e}_1, \lambda_1 \mathbf{e}_1, \cdots, \lambda_{k_1} \mathbf{e}_1, \mathbf{e}_2, \mu_1 \mathbf{e}_1 + \nu_1 \mathbf{e}_2, \cdots, \\ &\qquad \mu_{k_2} \mathbf{e}_1 + \nu_{k_2} \mathbf{e}_2, \mathbf{a}''_{i_3}, \cdots, \mathbf{a}''_N) \end{aligned}$$

となっている．ここで第 3 行と第 $j$ 行 ($j \geq 4$) を入れ替える，第 3 行の定数倍を他の行に加える，第 3 行を定数倍する，という行に関する基本変形をして

$$\begin{aligned} A''' = (&\mathbf{0}, \cdots, \mathbf{0}, \mathbf{e}_1, \lambda_1 \mathbf{e}_1, \cdots, \lambda_{k_1} \mathbf{e}_1, \mathbf{e}_2, \mu_1 \mathbf{e}_1 + \nu_1 \mathbf{e}_2, \cdots, \\ &\qquad \mu_{k_2} \mathbf{e}_1 + \nu_{k_2} \mathbf{e}_2, \mathbf{e}_3, \mathbf{a}'''_{i_3+1}, \cdots, \mathbf{a}'''_N) \end{aligned}$$

となる．この操作を続ければ $A$ は行に関する基本変形で次のような形にできる．

$$\begin{aligned} &E_k \cdots E_1 A \\ &= \begin{pmatrix} & 1 & * \cdots & * & 0 & * \cdots & * & 0 & * & \cdots\cdots * & 0 & * & \cdots \\ & 0 & \cdots & 0 & 1 & * \cdots & * & 0 & * & \cdots\cdots * & 0 & * & \cdots \\ & \vdots & & & 0 & \cdots & 0 & 1 & * & \cdots\cdots * & 0 & * & \cdots \\ & & & & \vdots & & & 0 & \ddots & \cdots\cdots * & 0 & * & \cdots \\ O & & & & & & & \vdots & & & \vdots & \vdots & \\ & & & & & O & & & & & 0 & * & \cdots \\ & & & & & & & & & & 1 & * & \cdots \\ & & & & & & & & & & 0 & \cdots & 0 \\ & & & & & & & & & & & O & \end{pmatrix} \end{aligned} \tag{11.2}$$

ここで $*$ はある数を表している．列の置換により，これは

$$\begin{pmatrix} I_{r'} & K \\ O & O \end{pmatrix}$$

と変形できる．したがって，この行列を $\Sigma$ とすると，有限個の $M$ 次基本行列

$E_1, \cdots, E_k$ と置換行列 $P$ が存在し,

$$E_k \cdots E_1 AP = \Sigma$$

である. $E_1, \cdots, E_k, P$ は正則行列であるから,

$$r = \mathrm{rank}(A) = \mathrm{rank}(E_k \cdots E_1 AP) = \mathrm{rank}(\Sigma)$$

である. $\Sigma$ のピボットの数は $r'$ であるから, 定理 11.14 より $\Sigma$ の階数は $r'$ である. したがって $r' = r$ となる. ■

**系 11. 18** $A$ を $(M, N)$ 行列とし, $r = \mathrm{rank}(A)$ とする. このとき, $M$ 次正則行列 $P$ と $N$ 次正則行列 $Q$ が存在し,

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

とできる.

**証明.** $\begin{pmatrix} I_r & K \\ O & O \end{pmatrix}$ の第 $j$ 列 $(1 \leq j \leq r)$ のスカラー倍を第 $k$ 列 $(r+1 \leq k \leq N)$ に加えることにより, $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ とすることができる. したがって, この操作の基本行列を $F_1, \cdots, F_n$ とすれば,

$$E_k \cdots E_1 APF_1 \cdots F_n = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

とできる. ■

定理 11.17 の証明より次のことがわかる.

**定理 11. 19** 任意の $(M, N)$ 行列 $A$ に対して, ある $M$ 次基本行列 $E_1, \cdots, E_k$ により, $E_1 \cdots E_k A$ が行既約階段行列となるようにできる.

# 第12章　連立1次方程式の一般解

本章では，次の一般的な連立 1 次方程式

$$\begin{cases} a_{11}x_1 + \cdots + a_{1N}x_N = b_1 \\ \qquad\qquad \vdots \\ a_{M1}x_1 + \cdots + a_{MN}x_N = b_M \end{cases} \tag{12.1}$$

の解の存在，特殊解，一般解などについて学ぶ．一般に $M > N$ の場合が方程式の数が未知数の数より多く，過剰決定系といい，$N > M$ の場合は方程式の数が未知数の数より少なく，不足決定系という．

**ポイント 12**　連立 1 次方程式の一般解とは何か．一般解を理解し，かつ具体的な連立 1 次方程式に対して一般解を求めることができるようにしてほしい．

## 12.1　解の存在と一般解

方程式 (12.1) を行列を用いて記述する．

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} = (\mathbf{a}_1, \cdots, \mathbf{a}_N), \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix} \in \boldsymbol{K}^M \tag{12.2}$$

とおくと，方程式 (12.1) は

$$A\mathbf{x} = \mathbf{b} \tag{12.3}$$

と表せる．方程式 (12.1) ないしは (12.3) が解をもつとき，この方程式は**可解**であるという．

さて，$A\mathbf{x} = \mathbf{b}$ が可解であるとき，解はただ一つの場合もあるが，複数存在す

ることもある．$A\mathbf{x}=\mathbf{b}$ の個々の解をこの方程式の**特殊解**という．一つの特殊解 $\mathbf{x}^{(0)}$ をとり固定する．このとき，任意の $\mathbf{u}\in\mathcal{N}(A)$ と $\mathbf{x}^{(0)}$ との和も $A\mathbf{x}=\mathbf{b}$ の解である．実際，

$$A(\mathbf{x}^{(0)}+\mathbf{u})=A\mathbf{x}^{(0)}+A\mathbf{u}=\mathbf{b}+\mathbf{0}=\mathbf{b}$$

となっている．じつは $A\mathbf{x}=\mathbf{b}$ の任意の解は $\mathbf{x}^{(0)}$ とある $\mathbf{u}\in\mathcal{N}(A)$ との和で表されている．このことを証明しよう．

**定理 12.1**　$A$ を $(M,N)$ 行列とし，$\mathbf{b}\in\boldsymbol{K}^M$ とする．$A\mathbf{x}=\mathbf{b}$ の解が存在するものとする．$A\mathbf{x}=\mathbf{b}$ の特殊解 $\mathbf{x}^{(0)}$ を一つとり，固定する．このとき，$A\mathbf{x}=\mathbf{b}$ の任意の解 $\mathbf{x}$ に対して，ある $\mathbf{u}\in\mathcal{N}(A)$ が存在し，

$$\mathbf{x}=\mathbf{x}^{(0)}+\mathbf{u}$$

と表せる．

**証明.** $\mathbf{u}=\mathbf{x}-\mathbf{x}^{(0)}$ とおく．このとき，

$$A\mathbf{u}=A(\mathbf{x}-\mathbf{x}^{(0)})=A\mathbf{x}-A\mathbf{x}^{(0)}=\mathbf{b}-\mathbf{b}=\mathbf{0}$$

であるから，$\mathbf{u}\in\mathcal{N}(A)$ である．これが求める $\mathbf{u}$ になっている．■

**定義 12.1**　$A$ を $(M,N)$ 行列とする．$\dim\mathcal{N}(A)\geq 1$ のとき，$\mathcal{N}(A)$ の基底 $\{\mathbf{u}_1,\cdots,\mathbf{u}_s\}$ $(s=\dim\mathcal{N}(A))$ を方程式 $A\mathbf{x}=\mathbf{0}$ の**基本解**という．また，$s=\dim\mathcal{N}(A)$ を方程式 $A\mathbf{x}=\mathbf{0}$ の**自由度**という．

定理 12.1 より次のことがわかる．

**定理 12.2**　$A$ を $(M,N)$ 行列とし，$s=\dim\mathcal{N}(A)$ とおく．$\mathbf{b}\in\boldsymbol{K}^M$ とし，$A\mathbf{x}=\mathbf{b}$ の解 $\mathbf{x}^{(0)}\in\boldsymbol{K}^N$ が存在するものとする．このとき次のことが成り立つ．

(1)　$\dim\mathcal{N}(A)=0$ の場合，$\mathbf{x}^{(0)}$ が $A\mathbf{x}=\mathbf{b}$ のただ一つの解である．

(2)　$\dim\mathcal{N}(A)\geq 1$ の場合，$\mathbf{u}_1,\cdots,\mathbf{u}_s\in\boldsymbol{K}^N$ を $A\mathbf{x}=\mathbf{0}$ の基本解とすると，任意の $\alpha_j\in\boldsymbol{K}$ $(j=1,\cdots,s)$ に対して

$$\mathbf{x}=\mathbf{x}^{(0)}+\alpha_1\mathbf{u}_1+\cdots+\alpha_s\mathbf{u}_s \tag{12.4}$$

は $A\mathbf{x}=\mathbf{b}$ の解である．逆に $A\mathbf{x}=\mathbf{b}$ の任意の解はある $\alpha_j\in\boldsymbol{K}$ $(j=1,\cdots,s)$ に対して (12.4) と表せる．(12.4) を $A\mathbf{x}=\mathbf{b}$ の**一般解**という．

**証明.** (1)　$\mathcal{N}(A)=\{\mathbf{0}\}$ であるから，定理 12.1 より，$A\mathbf{x}=\mathbf{b}$ の任意の解は $\mathbf{x}=\mathbf{x}^{(0)}+\mathbf{0}=\mathbf{x}^{(0)}$ である．

(2)　$\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ が $\mathcal{N}(A)$ の基底であることと，定理 12.1 およびその直前に示したことより明らか．■

どのような場合に方程式 (12.1) は解をもつのだろうか？　そしてどのような場合に解の一意性が成り立つだろうか？　これについての一つの判定法を導いておく．$A, \mathbf{b}$ は (12.2) で定めたものとする．このとき，行列

$$\widetilde{A} = (\mathbf{a}_1, \cdots, \mathbf{a}_N, \mathbf{b})$$

を拡大係数行列という．第 4 章では拡大係数行列を，方程式の係数と与えられた $\mathbf{b}$ とを区別するため

$$(\mathbf{a}_1, \cdots, \mathbf{a}_N \mid \mathbf{b})$$

と仕切りの縦棒を入れて記述していたが，本節では仕切りの縦棒は省略して行列として扱う．

$\mathbf{x} \in \boldsymbol{K}^N$ に対して，$\widetilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix} \in \boldsymbol{K}^{N+1}$ とすると，

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow \widetilde{A}\widetilde{\mathbf{x}} = 0$$

が成り立っている．明らかに $\widetilde{\mathbf{x}} \neq \mathbf{0}_{N+1}$ である．

**定理 12.3**　$A$ を $(M, N)$ 行列とし，$\mathbf{b} \in \boldsymbol{K}^M$ とする．このとき次の (1), (2) は同値である．

(1)　$A\mathbf{x} = \mathbf{b}$ をみたす $\mathbf{x} \in \boldsymbol{K}^N$ が存在する．

(2)　$\mathrm{rank}(A) = \mathrm{rank}(\widetilde{A})$.

**証明.** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とおく．$\mathbf{x} = (x_j) \in \boldsymbol{K}^N$ に対して

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_N\mathbf{a}_N$$

であることに注意すると，

$$\begin{aligned}(1) &\Longleftrightarrow \mathbf{b} \in \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} \\ &\Longleftrightarrow \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} = \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N, \mathbf{b}\} \\ &\Longleftrightarrow \dim \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N\} = \dim \mathrm{span}\,\{\mathbf{a}_1, \cdots, \mathbf{a}_N, \mathbf{b}\}\end{aligned}$$

したがって定理 11.1 より (1) ⇔ (2) が得られる．■

さらに，解が一意であるための必要十分条件として次のものがある．

**定理 12.4**　$A$ を $(M,N)$ 行列とし，$\mathbf{b}\in\boldsymbol{K}^M$ とする．このとき次の (1), (2) は同値である．

(1)　$A\mathbf{x}=\mathbf{b}$ がただ一つの解をもつ．

(2)　$\mathrm{rank}(A)=\mathrm{rank}(\widetilde{A})=N$.

**証明.** (1) を仮定する．このとき $\mathrm{rank}(A)=N$ であることを示せば前定理から (2) が得られる．もし $\mathrm{rank}(A)<N$ であるならば，定理 10.26 より $\dim\mathcal{N}(A)=N-\mathrm{rank}(A)\geq 1$ であるから，定理 12.2 (2) より $A\mathbf{x}=\mathbf{b}$ の解は一意的ではない．

(2) を仮定する．このとき前定理から $A\mathbf{x}=\mathbf{b}$ の解は少なくとも一つ存在する．また定理 10.26 より $\dim\mathcal{N}(A)=N-\mathrm{rank}(A)=0$ であるから，定理 12.2 (1) より解の一意性が成り立つ．■

**例題 12.1**　次の連立 1 次方程式の解の存在，一意性を判定せよ．

(1) $\begin{cases} x_2+2x_3=1 \\ x_1+3x_3=2 \\ 2x_1+3x_2=1 \end{cases}$　　(2) $\begin{cases} x_2+2x_3=1 \\ -x_1+3x_3=2 \\ -2x_1-3x_2=1 \end{cases}$

(3) $\begin{cases} x_2+2x_3=1 \\ -x_1+3x_3=1 \\ -2x_1-3x_2=1 \end{cases}$

**答.** (1)

$$A=\begin{pmatrix}0&1&2\\1&0&3\\2&3&0\end{pmatrix},\quad \widetilde{A}=\begin{pmatrix}0&1&2&1\\1&0&3&2\\2&3&0&1\end{pmatrix}$$

とする．$\widetilde{A}$ の行に関する基本変形をすると，

$$\begin{pmatrix}0&1&2&1\\1&0&3&2\\2&3&0&1\end{pmatrix}\xrightarrow[R_1\leftrightarrow R_2]{}\begin{pmatrix}1&0&3&2\\0&1&2&1\\2&3&0&1\end{pmatrix}\xrightarrow[R_3-2R_1]{}\begin{pmatrix}1&0&3&2\\0&1&2&1\\0&3&-6&-3\end{pmatrix}$$

$$\xrightarrow[R_3-3R_2]{}\begin{pmatrix}1&0&3&2\\0&1&2&1\\0&0&-12&-6\end{pmatrix}\xrightarrow[-\frac{1}{12}R_3]{}\begin{pmatrix}1&0&3&2\\0&1&2&1\\0&0&1&\frac{1}{2}\end{pmatrix}.$$

また同じ変形で $A$ は

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

になる．したがって，

$$\mathrm{rank}(A) = \mathrm{rank}(\widetilde{A}) = 3$$

であるから，解はただ一つ存在する．

(2)

$$A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 0 & 3 & 2 \\ -2 & -3 & 0 & 1 \end{pmatrix}$$

とする．行に関する基本変形で次のようにする．

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 0 & 3 & 2 \\ -2 & -3 & 0 & 1 \end{pmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{pmatrix} -1 & 0 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[-R_1]{} \begin{pmatrix} 1 & 0 & -3 & -2 \\ 0 & 1 & 2 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix} \xrightarrow[R_3+2R_1]{} \begin{pmatrix} 1 & 0 & -3 & -2 \\ 0 & 1 & 2 & 1 \\ 0 & -3 & -6 & -3 \end{pmatrix}$$

$$\xrightarrow[R_3+3R_2]{} \begin{pmatrix} 1 & 0 & -3 & -2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

同様の基本変形により $A$ は $\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ であるから，

$$\mathrm{rank}(A) = \mathrm{rank}(\widetilde{A}) = 2 < 3$$

である．したがって，解は存在するが，一意的ではないことがわかる．実際，$x_3 = s$ として

$$\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ s \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

を解くと，$x_1 - 3s = -2$, $x_2 + 2s = 1$ より

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2+3s \\ 1-2s \\ s \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

はすべての $s \in \boldsymbol{K}$ に対して連立 1 次方程式 (2) の解になっている．

(3)

$$A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 0 & 3 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix}$$

とする．行に関する基本変形により

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 0 & 3 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{pmatrix} -1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[-R_1]{} \begin{pmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & 2 & 1 \\ -2 & -3 & 0 & 1 \end{pmatrix} \xrightarrow[R_3+2R_1]{} \begin{pmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & -3 & -6 & -1 \end{pmatrix}$$

$$\xrightarrow[R_3+3R_2]{} \begin{pmatrix} 1 & 0 & -3 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

同様の変形で $A$ は $\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ となるので，

$$\mathrm{rank}(A) = 2 \neq 3 = \mathrm{rank}(\widetilde{A})$$

である．ゆえに解は存在しない． ■

## 12.2 連立 1 次方程式の一般解の求め方

具体的に連立 1 次方程式の特殊解，一般解を求める方法を述べる．はじめに

$$Q = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix} \tag{12.5}$$

の形の行列を係数行列とする方程式の解について考える．すでに示したように，任意の行列 $A \in \boldsymbol{K}^{M \times N}$ は行に関するある有限回の基本変形 $E_1, \cdots, E_k$ と列の置換 $P$ により

$$E_k \cdots E_1 AP = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix} \tag{12.6}$$

の形に変形可能である．まず方程式 $Q\mathbf{y} = \mathbf{c}$ の解について論じた後，(12.6) を用いて，方程式 $A\mathbf{x} = \mathbf{b}$ の解について調べる．

**定理 12.5**　$Q$ を $(M, N)$ 行列で，$Q = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix}$ とする．ただし $r < M$ とする．このとき，$\mathbf{c} = (c_j)_{j=1,\cdots,M} \in \boldsymbol{K}^M$ に対して

$$Q\mathbf{y} = \mathbf{c} \tag{12.7}$$

が解をもつための必要十分条件は，

$$c_{r+1} = \cdots = c_M = 0 \tag{12.8}$$

が成り立つことである．この場合には，$\mathbf{c}$ は方程式 (12.7) の特殊解であり，$K = (k_{ij})_{\substack{i=1,\cdots,r \\ j=1,\cdots,N-r}}$ とおくと，

$$\mathbf{u}_1 = \begin{pmatrix} -k_{11} \\ \vdots \\ -k_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad \mathbf{u}_{N-r} = \begin{pmatrix} -k_{1,N-r} \\ \vdots \\ -k_{r,N-r} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{12.9}$$

は $Q\mathbf{y} = 0$ の基本解である．したがって (12.7) の一般解は

$$\mathbf{y} = \mathbf{c} + s_1\mathbf{u}_1 + \cdots + s_{N-r}\mathbf{u}_{N-r}$$

である．

**証明．** もしある $r+1 \leq i \leq M$ に対して，$c_i \neq 0$ ならば，明らかに $\mathrm{rank}(Q, \mathbf{c}) = r+1 > \mathrm{rank}(Q)$ であるから，$A\mathbf{y} = \mathbf{c}$ の解は存在しない．したがって，$Q\mathbf{y} = \mathbf{c}$ の解が存在すれば $c_i = 0\,(i = r+1, \cdots, M)$ でなければならない．逆に $c_i = 0\,(i = r+1, \cdots, M)$ であれば，$\mathrm{rank}(Q, \mathbf{c}) = r = \mathrm{rank}(Q)$ であるから，$Q\mathbf{y} = \mathbf{c}$ の解

が存在する．これで前半の証明がなされた．

後半の主張を証明する．$\mathbf{y} = (y_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ に対して

$$Qy = \begin{pmatrix} y_1 + y_{r+1}k_{11} + \cdots + y_N k_{1,N-r} \\ \vdots \\ y_r + y_{r+1}k_{r1} + \cdots + y_N k_{r,N-r} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

であるから，$Q\mathbf{y} = \mathbf{0}$ が成り立つための必要十分条件は

$$\begin{cases} y_1 + y_{r+1}k_{11} + \cdots + y_N k_{1,N-r} = 0 \\ \qquad\qquad \vdots \\ y_r + y_{r+1}k_{r1} + \cdots + y_N k_{r,N-r} = 0 \end{cases} \tag{12.10}$$

である．いま $Q\mathbf{y} = \mathbf{0}$ とし，$s_1, \cdots, s_{N-r} \in \boldsymbol{K}$ とする．$y_{r+1} = s_1, \cdots, y_N = s_{N-r}$ とおくと，(12.10) より

$$\begin{aligned} y_1 &= s_1(-k_{11}) + \cdots + s_{N-r}(-k_{1,N-r}) \\ &\vdots \\ y_r &= s_1(-k_{r1}) + \cdots + s_{N-r}(-k_{r,N-r}) \\ y_{r+1} &= s_1 \\ &\vdots \\ y_N &= s_{N-r} \end{aligned}$$

となっている．したがって，任意の $s_1, \cdots, s_{N-r} \in \boldsymbol{K}$ に対して

$$\mathbf{y} = s_1\mathbf{u}_1 + \cdots + s_{N-r}\mathbf{u}_{N-r}$$

と表せている．ゆえに $\mathcal{N}(Q) \subset \operatorname{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_{N-r}\}$ が得られる．

容易に $\mathbf{u}_1, \cdots, \mathbf{u}_{N-r} \in \mathcal{N}(Q)$ であることがわかるので，

$$\mathcal{N}(Q) = \operatorname{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_{N-r}\}$$

であり，また明らかに $\mathbf{u}_1, \cdots, \mathbf{u}_{N-r}$ は線形独立であるから，$\mathbf{u}_1, \cdots, \mathbf{u}_{N-r}$ は $Q\mathbf{y} = \mathbf{0}$ の基本解である．また $c_{r+1} = c_N = 0$ の場合を考えているので，$Q\mathbf{c} = \mathbf{c}$ である．よって定理が証明された．■

一般の行列の場合は，$Q$ の形に標準化することにより次のように解を求めることができる．

**定理 12.6**　$A$ を $(M,N)$ 行列で，$r=\mathrm{rank}(A)$ とする．$A$ を行に関する基本変形 $E_1,\cdots,E_k$ と置換行列 $P$ により (12.6) となるようにする．(12.6) の右辺の行列を $Q$ とおく．このとき次のことが成り立つ．

(1)　(12.9) で定めた $Q\mathbf{x}=\mathbf{0}$ の基本解 $\mathbf{u}_1,\cdots,\mathbf{u}_{N-r}$ に対して，$P\mathbf{u}_1,\cdots,P\mathbf{u}_{N-r}$ は $A\mathbf{x}=\mathbf{0}$ の基本解である．

(2)　$\mathbf{b}\in\boldsymbol{K}^M$ とする．方程式 $A\mathbf{x}=\mathbf{b}$ が解をもつための必要十分条件は

$$E_k\cdots E_1\mathbf{b}\,(j)=0,\quad j=r+1,\cdots,M \tag{12.11}$$

が成り立つことである (ここで $E_k\cdots E_1\mathbf{b}\,(j)$ は $E_k\cdots E_1\mathbf{b}$ の第 $j$ 成分を表している)．この場合，$PE_k\cdots E_1\mathbf{b}$ が方程式 $A\mathbf{x}=\mathbf{b}$ の特殊解になっている．

(3)　また (12.11) が成り立つとき，

$$PE_k\cdots E_1\mathbf{b}+\sum_{j=1}^{N-r}s_jP\mathbf{u}_j,\quad s_1,\cdots,s_{N-r}\in\boldsymbol{K}$$

は $A\mathbf{x}=\mathbf{b}$ の一般解である．

**証明.** (1)　$j=1,\cdots,N-r$ に対して $E_k\cdots E_1AP\mathbf{u}_j=Q\mathbf{u}_j=\mathbf{0}$ であり，$E_k\cdots E_1$ は正則行列であるから，$AP\mathbf{u}_j=\mathbf{0}$ である．$\mathbf{x}'$ を $A\mathbf{x}=\mathbf{0}$ の任意の解とする．このとき，$E_k\cdots E_1APP^{-1}\mathbf{x}'=\mathbf{0}$ であるから，定理 12.5 より，$P^{-1}\mathbf{x}'=\sum_{j=1}^{N-r}\lambda_j\mathbf{u}_j$ と表せる．ゆえに $\mathbf{x}'=\sum_{j=1}^{N-r}\lambda_jP\mathbf{u}_j$ である．また $P\mathbf{u}_1,\cdots,P\mathbf{u}_{N-r}$ は線形独立であるから，$\mathcal{N}(A)$ の基底，すなわち $A\mathbf{x}=\mathbf{0}$ の基本解になっている．

(2)　$A\mathbf{x}=\mathbf{b}$ が解をもつならば，解 $\mathbf{x}$ に対して

$$QP^{-1}\mathbf{x}=E_k\cdots E_1APP^{-1}\mathbf{x}=E_k\cdots E_1\mathbf{b}$$

が成り立つ．ゆえに定理 12.5，(12.8) より (12.11) が成立する．逆に (12.11) が成り立てば，定理 12.5 より $Q\,(E_k\cdots E_1\mathbf{b})=E_k\cdots E_1\mathbf{b}$ が成り立つ．そこで

$$\mathbf{x}^{(0)}=P\,(E_k\cdots E_1\mathbf{b})$$

とおけば，

$$\begin{aligned}A\mathbf{x}^{(0)}&=E_1^{-1}\cdots E_k^{-1}QP^{-1}\mathbf{x}^{(0)}=E_1^{-1}\cdots E_k^{-1}QP^{-1}P\,(E_k\cdots E_1\mathbf{b})\\&=E_1^{-1}\cdots E_k^{-1}Q\,(E_k\cdots E_1\mathbf{b})=E_1^{-1}\cdots E_k^{-1}E_k\cdots E_1\mathbf{b}\\&=\mathbf{b}\end{aligned}$$

が得られる．

(1), (2) より (3) は明らかである． ■

### 一般解の求め方

これまでの定理の証明で得られたことをもとに，方程式の一般解の求め方をまとめておく．まず定理 12.5 の証明から，方程式 (12.7) の解は次のようにすれば求めることができる．

$$y_{r+1} = s_1, \quad y_{r+2} = s_2, \quad \cdots, \quad y_N = s_{N-r}$$

(ただし $s_1, \cdots, s_{N-r}$ は任意定数) とおき，

$$y_1 = -\sum_{j=1}^{N-r} k_{1j}s_j + c_1, \quad \cdots, \quad y_r = -\sum_{j=1}^{N-r} k_{rj}s_j + c_r$$

とおくと $\mathbf{y}^T = (\, y_1 \quad \cdots \quad y_N \,)$ が $Q\mathbf{y} = \mathbf{c}$ の一般解である．

方程式 $A\mathbf{x} = \mathbf{b}$ を解く場合は次のようにして解を求めることができる．係数行列 $A$ を行に関する基本変形により行既約階段行列

$$\left(\begin{array}{ccccccccccccc}
 & \boxed{1} & *\cdots & * & 0 & *\cdots & * & 0 & * & \cdots\cdots* & 0 & * & \cdots \\
 & 0 & \cdots & 0 & \boxed{1} & *\cdots & * & 0 & * & \cdots\cdots* & 0 & * & \cdots \\
 & \vdots & & & 0 & \cdots & 0 & \boxed{1} & * & \cdots\cdots* & 0 & * & \cdots \\
 & & & & \vdots & & & 0 & \ddots & \cdots\cdots* & 0 & * & \cdots \\
O & & & & & & & \vdots & & & \vdots & \vdots & \\
 & & & & & O & & & & & 0 & * & \cdots \\
 & & & & & & & & & & \boxed{1} & * & \cdots \\
 & & & & & & & & & & 0 & \cdots & 0 \\
 & & & & & & & & & & & O &
\end{array}\right)$$

に変形する．この行既約階段行列を $A' = (a'_{jk})$ とおき，$A$ を $A'$ に変形するのに行った基本変形と同じ基本変形により $\mathbf{b}$ が $\mathbf{b}' = (b'_j)$ になったとする．$A'$ のピボットを $a'_{j,k(j)}$, $j = 1, \cdots, r$ とおく．$A'$ を列に関する置換により標準形 $Q$ にしたので，そのことを考慮に入れて，

$$x_k = s_k, \quad 1 \leq k \leq N, \quad k \neq k(1), \cdots, k(r)$$

とおく．ただしここで $s_k$ は任意定数である．次に

$$x_{k(1)} = -\sum_{\substack{k \geq k(1) \\ k \neq k(1), \cdots, k(r)}} a'_{1,k} s_k + b'_{k(1)},$$

$$x_{k(2)} = -\sum_{\substack{k \geq k(2) \\ k \neq k(2), \cdots, k(r)}} a'_{1,k} s_k + b'_{k(2)},$$

$$\cdots\cdots$$

$$x_{k(r)} = -\sum_{\substack{k \geq k(r) \\ k \neq k(r)}} a'_{1,k} s_k + b'_{k(r)}$$

とおく．このとき $N-r$ 個の任意定数 $s_k$ を含んだ $(x_1 \quad \cdots \quad x_N)$ が方程式 $A\mathbf{x} = \mathbf{b}$ の解である．

**例題 12.2**　次の方程式の一般解を求めよ．

$$\begin{pmatrix} 0 & 0 & 2 & 1 & 1 \\ 1 & -1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 8 \\ 0 \end{pmatrix}.$$

**答．** 拡大係数行列を行に関する基本変形により行既約階段行列に変形する．

$$\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & -1 & 2 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & -1 & 2 & 3 & 4 & 3 \\ 0 & 0 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccccc|c} 1 & -1 & 2 & 3 & 4 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & -1 & 0 & 2 & 3 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccccc|c} \boxed{1} & -1 & 0 & 2 & 0 & -3 \\ 0 & 0 & \boxed{1} & \frac{1}{2} & 0 & -1 \\ 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

$x_2 = s,\ x_4 = t$ を任意定数とし，

$$\begin{aligned} x_1 &= s - 2t - 3, \\ x_3 &= -\frac{1}{2}t - 1, \\ x_5 &= 2 \end{aligned}$$

とする．このとき

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} s - 2t - 3 \\ s \\ -\frac{1}{2}t - 1 \\ t \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ -1 \\ 0 \\ 2 \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$$

が一般解である．■

**注意 12.1（別解法）** 例題 12.2 の解法より

$$\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & -1 & 2 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccccc|c} 1 & -1 & 0 & 2 & 0 & -3 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

と変形されるが，さらにこれは列による置換で

$$\underset{\substack{\begin{pmatrix} 1\,2\,3\,4\,5 \\ 1\,4\,2\,5\,3 \end{pmatrix} \\ \text{により列を置換}}}{\longrightarrow} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & -1 & 2 & -3 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & -1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

と標準化される．ここで行った列に関する置換 $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix}$ を表現する置換行列を $P$ とする．すなわち

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

とする．定理 12.5 より，変形後の方程式の一般解は

$$\begin{pmatrix} -3 \\ -1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + s_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} -2 \\ -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

となる．したがって，これに $P$ を右から施して得られた

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ -1 \\ 0 \\ 2 \end{pmatrix} + s_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} -2 \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$$

が求める一般解である (注：最後の変形は $\sigma$ の下の数に上の数を対応させて，第 1 列 → 第 1 列，第 4 列 → 第 2 列，第 2 列 → 第 3 列，第 5 列 → 第 4 列，第 3 列 → 第 5 列の入れ替えを行うと覚えておくと記憶しやすい).

**例題 12.3**　次の方程式の一般解を求めよ．

$$\begin{cases} x_2 + 2x_3 - x_5 = 4 \\ 2x_4 - x_5 = 4 \\ 2x_5 = 4 \end{cases}$$

**答．** 行に関する基本変形により，

$$\left(\begin{array}{ccccc|c} 0 & 1 & 2 & 0 & -1 & 4 \\ 0 & 0 & 0 & 2 & -1 & 4 \\ 0 & 0 & 0 & 0 & 2 & 4 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 0 & 1 & 2 & 0 & -1 & 4 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccccc|c} 0 & \boxed{1} & 2 & 0 & 0 & 6 \\ 0 & 0 & 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & 0 & 0 & \boxed{1} & 2 \end{array}\right).$$

したがって，$s, t$ を任意定数として，

$$
\begin{aligned}
&x_1 = s, \quad x_3 = t, \\
&x_5 = 2, \quad x_4 = 3, \\
&x_2 = -2t + 6,
\end{aligned}
$$

すなわち

$$
\begin{pmatrix} 0 \\ 6 \\ 0 \\ 3 \\ 2 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}
$$

が一般解である． ■

# 第13章　基底変換と行列の対角化

一般にものごとを別の視点と尺度を用いて眺めると見えなかった性質が見えてくることがある．たとえば幾何学の場合は座標系を変えることにより，今まで見えてこなかった対象の特性が見えてくることがある．ここでは座標系を変える，すなわち座標変換とその線形代数における効用について述べる．

**ポイント 13**　この章のポイントは，$\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像がある条件をみたすならば，$\boldsymbol{K}^N$ の標準座標系 (定義は後述) を座標変換して，新しい座標系で考えると，その表現行列が対角行列となって現れるようにできることを示す．この場合，一般には変換後の座標系が斜交座標になっていることもある．しかし線形写像によっては，変換後の座標系が直交座標となるようにできることもある．それはどのような場合かを調べる．本章のキーワードは，対角化と固有値である．

まず座標系の定義からはじめる．$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $\boldsymbol{K}^N$ の基底とする．このとき，組 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $\boldsymbol{K}^N$ の**座標系**という．特に $\mathcal{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_N\}$ を $\boldsymbol{K}^N$ の**標準座標系**という．$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の正規直交基底であるとき，$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を**直交座標系**といい，そうでない座標系を**斜交座標系**という．標準座標系では，$\mathbf{x} = (x_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ は

$$\mathbf{x} = \sum_{j=1}^{N} x_j \mathbf{e}_j$$

と表せている．一方，別の座標系 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ に対しても

$$\mathbf{x} = \sum_{j=1}^{N} y_j \mathbf{u}_j$$

と一意的に表せる．ここで $(y_j)_{j=1,\cdots,N} \in \boldsymbol{K}^N$ を座標系 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ に関

する $\mathbf{x}$ の座標，あるいは $\mathbf{x}$ **の基底** $\mathcal{U}$ **による表示**，あるいは $\mathbf{x}$ **の** $\mathcal{U}$**-座標**という．そして

$$\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$$

と表す．$\mathbf{x}_{\mathcal{U}}$ を $(\mathbf{x})_{\mathcal{U}}$ と表すこともある．特に $\mathbf{x}$ の基底 $\mathcal{E}$ による表示を

$$\mathbf{x}_{\mathcal{E}} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \mathbf{x}$$

とみなすことができる．

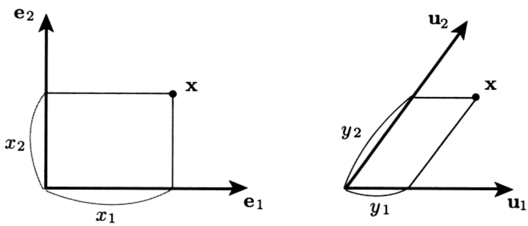



図 13.1

幾何学などでは座標変換すると問題が簡単になることがあるが，線形代数でも基底を標準基底でなく別の基底を用いて考えると問題が簡単になることがしばしばある．このことについて述べるため，$\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像 $T$ を考える．$T$ の表現行列を求めて見よう．

$$\mathbf{a}_j = T(\mathbf{e}_j), \quad j = 1, \cdots, N$$

とする．

$$\mathbf{a}_j = (\mathbf{a}_j)_{\mathcal{E}} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{Nj} \end{pmatrix}$$

と表し，

$$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N) = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$$

とおくと,

$$T(\mathbf{x}) = A\mathbf{x}$$

となっている．$A$ は線形写像 $T$ によるベクトルの動きを $\mathcal{E}$-座標のもとで眺めた結果得られた行列である．これを $\mathcal{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_N\}$ の代わりに別の基底 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を基に考えてみることにする．すなわち,

$$\mathbf{b}_j = T(\mathbf{u}_j), \quad j = 1, \cdots, N$$

とし,

$$(\mathbf{b}_j)_{\mathcal{U}} = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{Nj} \end{pmatrix}$$

とする．このとき

$$B = ((\mathbf{b}_1)_{\mathcal{U}}, \cdots, (\mathbf{b}_N)_{\mathcal{U}}) = \begin{pmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{pmatrix}$$

とすると,

$$T(\mathbf{x})_{\mathcal{U}} = B\mathbf{x}_{\mathcal{U}}$$

となっていることが次のようにして証明できる．

$\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$ とし, $T(\mathbf{x})_{\mathcal{U}} = \begin{pmatrix} z_1 \\ \vdots \\ z_N \end{pmatrix}$ とする．このとき, $\mathbf{x} = \sum_{j=1}^{N} y_j \mathbf{u}_j$, $T(\mathbf{x}) = \sum_{i=1}^{N} z_i \mathbf{u}_i$ であるから,

$$\sum_{i=1}^{N} z_i \mathbf{u}_i = T(\mathbf{x}) = \sum_{j=1}^{N} y_j T(\mathbf{u}_j) = \sum_{j=1}^{N} y_j \sum_{i=1}^{N} b_{ij} \mathbf{u}_i$$

$$= \sum_{i=1}^{N} \left( \sum_{j=1}^{N} b_{ij} y_j \right) \mathbf{u}_i.$$

したがって

$$z_i = \sum_{j=1}^{N} b_{ij} y_j.$$

これより

$$T(\mathbf{x})_{\mathcal{U}} = \begin{pmatrix} z_1 \\ \vdots \\ z_N \end{pmatrix} = B \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = B\mathbf{x}_{\mathcal{U}}.$$

以下では $B$ を $T_{\mathcal{U}}$ と表し，$T$ の $\mathcal{U}$ に関する**表示**あるいは $\mathcal{U}$ に関する**表現行列**という．以上のことを定理の形にまとめると次のようになる．

**定理 13.1** $T(\mathbf{u}_j) = \sum_{i=1}^{N} b_{ij}\mathbf{u}_i$ とし，

$$T_{\mathcal{U}} = \begin{pmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{pmatrix}$$

とおく．このとき

$$T(\mathbf{x})_{\mathcal{U}} = T_{\mathcal{U}}\mathbf{x}_{\mathcal{U}}$$

である．

$T_{\mathcal{U}}$ は基底 $\mathcal{U}$ のとり方によりさまざまな行列になる．この章では基底を適切に選ぶことにより $T_{\mathcal{U}}$ を簡単な形にできることを学ぶ．

## 13.1 基底変換

まず基底が変わると座標がどのように変わるかを考える．$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$, $\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_N\}$ を $\boldsymbol{K}^N$ の二組の基底とする．このとき，$j = 1, \cdots, N$ に対して

$$\mathbf{u}_j = \sum_{i=1}^{N} \alpha_{ij}\mathbf{v}_i$$

と表せる．すなわち，

$$(\mathbf{u}_j)_{\mathcal{V}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{Nj} \end{pmatrix} \tag{13.1}$$

である．したがって，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$\mathbf{x} = \sum_{j=1}^{N} a_j \mathbf{u}_j = \sum_{j=1}^{N} a_j \sum_{i=1}^{N} \alpha_{ij} \mathbf{v}_i = \sum_{i=1}^{N} \left( \sum_{j=1}^{N} \alpha_{ij} a_j \right) \mathbf{v}_i,$$

$$\mathbf{x} = \sum_{i=1}^{N} b_i \mathbf{v}_i$$

のとき，

$$b_i = \sum_{j=1}^{N} \alpha_{ij}\, a_j$$

が得られる．ゆえに

$$\begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1N} \\ \vdots & \ddots & \vdots \\ \alpha_{N1} & \cdots & \alpha_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \tag{13.2}$$

が成立する．以上の考察を定理の形にまとめておく．

**定理 13.2**　$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}, \mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_N\}$ を $\boldsymbol{K}^N$ の二組の基底とし，

$$\mathbf{u}_j = \sum_{i=1}^{N} \alpha_{ij} \mathbf{v}_i$$

とする．

$$X_{\mathcal{V} \leftarrow \mathcal{U}} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1N} \\ \vdots & \ddots & \vdots \\ \alpha_{N1} & \cdots & \alpha_{NN} \end{pmatrix} \tag{13.3}$$

とおく．このとき

$$\mathbf{x}_{\mathcal{V}} = X_{\mathcal{V} \leftarrow \mathcal{U}}\, \mathbf{x}_{\mathcal{U}}$$

が成り立つ．

**定義 13.1**　定理 13.2 で定義された $X_{\mathcal{V} \leftarrow \mathcal{U}}$ を $\mathcal{U}$ から $\mathcal{V}$ への**変換行列**という．

特に，

$$\mathbf{u}_j = (\mathbf{u}_j)_{\mathcal{E}} = \begin{pmatrix} u_{1j} \\ \vdots \\ u_{Nj} \end{pmatrix}, \quad j = 1, \cdots, N$$

のとき，

$$X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix} = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$$

である．

**定理 13.3**　$X_{\mathcal{V}\leftarrow\mathcal{U}}$ は正則行列であり，$X_{\mathcal{V}\leftarrow\mathcal{U}}^{-1} = X_{\mathcal{U}\leftarrow\mathcal{V}}$ である．

**証明.** $\mathcal{U}$ と $\mathcal{V}$ の役割を入れ替えると $X_{\mathcal{U}\leftarrow\mathcal{V}}$ が定義され，$\mathbf{x}_{\mathcal{U}} = X_{\mathcal{U}\leftarrow\mathcal{V}}\,\mathbf{x}_{\mathcal{V}}$ が任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して成立する．したがって，

$$\mathbf{x}_{\mathcal{U}} = X_{\mathcal{U}\leftarrow\mathcal{V}}\,\mathbf{x}_{\mathcal{V}} = X_{\mathcal{U}\leftarrow\mathcal{V}}\,X_{\mathcal{V}\leftarrow\mathcal{U}}\,\mathbf{x}_{\mathcal{U}}$$

が得られる．ゆえに，$X_{\mathcal{U}\leftarrow\mathcal{V}}\,X_{\mathcal{V}\leftarrow\mathcal{U}} = I_N$ となる．■

$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$, $\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_N\}$, $\mathcal{W} = \{\mathbf{w}_1, \cdots, \mathbf{w}_N\}$ を $\boldsymbol{K}^N$ の三組の基底とする．変換行列の間には次の関係式が成り立つ．

**定理 13.4**　$X_{\mathcal{W}\leftarrow\mathcal{U}} = X_{\mathcal{W}\leftarrow\mathcal{V}}X_{\mathcal{V}\leftarrow\mathcal{U}}$.

**証明.** $\mathbf{v}_i = \sum_{k=1}^{N} v_{ki}\mathbf{w}_k$ とすると，$X_{\mathcal{W}\leftarrow\mathcal{V}} = (v_{ij})$ であり，$\mathbf{u}_j = \sum_{k=1}^{N} w_{kj}\mathbf{w}_k$ とすると，$X_{\mathcal{W}\leftarrow\mathcal{U}} = (w_{ij})$ である．$X_{\mathcal{V}\leftarrow\mathcal{U}} = (u_{ij})$ とおくと

$$\mathbf{u}_j = \sum_{i=1}^{N} u_{ij}\mathbf{v}_i = \sum_{i=1}^{N} u_{ij} \sum_{k=1}^{N} v_{ki}\mathbf{w}_k = \sum_{k=1}^{N}\left(\sum_{i=1}^{N} v_{ki}u_{ij}\right)\mathbf{w}_k$$

であるから，

$$w_{kj} = \sum_{i=1}^{N} v_{ki}u_{ij}$$

となる．よって

$$X_{\mathcal{W}\leftarrow\mathcal{U}} = X_{\mathcal{W}\leftarrow\mathcal{V}}X_{\mathcal{V}\leftarrow\mathcal{U}}$$

が得られる．■

**例 13.1**　$\mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ とする．行に関する基本変形により

$$\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 1 \\ 0 & \boxed{3} \end{pmatrix}$$

となる．ピボットは 2 個あるので，$\mathrm{rank}\,(\mathbf{u}_1, \mathbf{u}_2) = 2$, すなわち $\mathbf{u}_1, \mathbf{u}_2$ は基底である．

$$X_{\mathcal{E}\leftarrow\mathcal{U}} = (\mathbf{u}_1, \mathbf{u}_2) = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$$

である．これの逆行列は，

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 3 & 1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \end{array}\right).$$

したがって，

$$X_{\mathcal{U}\leftarrow\mathcal{E}} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$$

である．これより，たとえば $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $(= \mathbf{x}_{\mathcal{E}})$ の $\mathcal{U}$ に関する座標は

$$\mathbf{x}_{\mathcal{U}} = X_{\mathcal{U}\leftarrow\mathcal{E}}\,\mathbf{x} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

である．

**例題 13.1**　$\mathcal{U} = \left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$, $\mathcal{V} = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}\right\}$ とする．$\mathcal{U}$ も $\mathcal{V}$ も $\boldsymbol{K}^2$ の基底であることを示せ．また，変換行列 $X_{\mathcal{V}\leftarrow\mathcal{U}}$ を求めよ．

**答．** $\mathcal{U}$ が基底であることは例 13.1 で示した．$\mathcal{V}$ については

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 3 & -2 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

より次のことがわかる．$\mathrm{rank}\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = 2$ であり，したがって $\mathcal{V}$ は基底であり，さらに

$$X_{\mathcal{V}\leftarrow\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}.$$

ゆえに

$$X_{\mathcal{V}\leftarrow\mathcal{U}} = X_{\mathcal{V}\leftarrow\mathcal{E}} X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -2 & 1 \end{pmatrix}.$$

■

**問題 13.1**

$$\mathcal{U} = \left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \mathcal{V} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

がそれぞれ $\boldsymbol{K}^3$ の基底であることを示し，変換行列 $X_{\mathcal{V}\leftarrow\mathcal{U}}$ を求めよ．$\mathbf{x}_{\mathcal{U}}^T = (1 \quad 0 \quad 0)$ に対して $\mathbf{x}_{\mathcal{V}}$ および $\mathbf{x}_{\mathcal{E}}$ を求めよ．

$\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像を $T$ とする．$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$, $\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ を $\boldsymbol{K}^N$ の基底とする．このとき $T_{\mathcal{U}}, T_{\mathcal{V}}$ の間には次の変換公式が成り立つ．

**定理 13.5**

$$T_{\mathcal{U}} = X_{\mathcal{U}\leftarrow\mathcal{V}}\, T_{\mathcal{V}}\, X_{\mathcal{V}\leftarrow\mathcal{U}}.$$

**証明.** すでに示したように

$$(T\mathbf{x})_{\mathcal{V}} = T_{\mathcal{V}}\mathbf{x}_{\mathcal{V}} = T_{\mathcal{V}} X_{\mathcal{V}\leftarrow\mathcal{U}}\, \mathbf{x}_{\mathcal{U}}.$$

一方，

$$(T\mathbf{x})_{\mathcal{V}} = X_{\mathcal{V}\leftarrow\mathcal{U}}(T\mathbf{x})_{\mathcal{U}} = X_{\mathcal{V}\leftarrow\mathcal{U}}\, T_{\mathcal{U}}\, \mathbf{x}_{\mathcal{U}}$$

であるから，

$$T_{\mathcal{V}}\, X_{\mathcal{V}\leftarrow\mathcal{U}}\, \mathbf{x}_{\mathcal{U}} = X_{\mathcal{V}\leftarrow\mathcal{U}}\, T_{\mathcal{U}}\, \mathbf{x}_{\mathcal{U}}.$$

ゆえに，

$$X_{\mathcal{V}\leftarrow\mathcal{U}}^{-1}\, T_{\mathcal{V}}\, X_{\mathcal{V}\leftarrow\mathcal{U}}\, \mathbf{x}_{\mathcal{U}} = T_{\mathcal{U}}\, \mathbf{x}_{\mathcal{U}}$$

である．よって，

$$T_{\mathcal{U}} = X_{\mathcal{U}\leftarrow\mathcal{V}}\, T_{\mathcal{V}}\, X_{\mathcal{V}\leftarrow\mathcal{U}}$$

が成立する．■

## 13.2　対角化と固有値

行列

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$$

が与えられたとき,

$$T\mathbf{x} = A\mathbf{x}$$

により $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像 $T$ を定める．このとき $T$ の $\mathcal{E}$-座標における表現行列は

$$T_{\mathcal{E}} = A$$

である．$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ が $\boldsymbol{K}^N$ の基底であるとき，$T$ の $\mathcal{U}$-座標における表現行列 $T_{\mathcal{U}}$ を $A_{\mathcal{U}}$ と表す．変換公式より $T_{\mathcal{U}} = X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1}\, T_{\mathcal{E}}\, X_{\mathcal{E}\leftarrow\mathcal{U}}$ であるから,

$$A_{\mathcal{U}} = X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1}\, A\, X_{\mathcal{E}\leftarrow\mathcal{U}}$$

が成立する．

行列 $A$ によっては，基底 $\mathcal{U}$ をうまくとることにより，$A_{\mathcal{U}}$ が対角行列になるように変形することが可能である．たとえば次のような例がある．

**例 13.2**　$T_{\mathcal{E}} = A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ とする．$\mathcal{U} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ とすると

$$X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

したがって,

$$A_{\mathcal{U}} = X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1} A X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$
$$= \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

以下，本章では主に次の問題を考える．

**課題**　$N$ 次正方行列 $A$ が与えられたとき，基底 $\mathcal{U}$ を選んで,

$$X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1}\,A\,X_{\mathcal{E}\leftarrow\mathcal{U}} = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)\,. \tag{13.4}$$

($\mathrm{diag}(\lambda_1, \cdots, \lambda_N)$ の定義は (2.8) 参照) とできるための条件は何か？
また (13.4) となるような基底 $\mathcal{U}$ はどのように見出せるのだろうか？
このようなことができるための行列 $A$ の条件は何か？

**定義 13.2** 基底 $\mathcal{U}$ を適切に選んで，行列 $A$ を (13.4) のように対角行列に変形できるとき，行列 $A$ は**対角化可能**であるという．

すでに示したように，$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が基底であるための必要十分条件は，行列 $(\mathbf{u}_1, \cdots, \mathbf{u}_N)$ が正則行列になることである．したがって，$A$ が対角化可能であるとは，ある正則行列 $P$ で

$$P^{-1}AP = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N) \tag{13.5}$$

が成り立つことと言い換えることができる．

ところで，行列を対角化したときに現れる対角成分は何であろうか？　結論を先に言えば，それは行列の固有値と呼ばれる量になっている．ここで行列の固有値と固有ベクトルの定義をしておきたい．

**定義 13.3** $A$ を $N$ 次正方行列とする．$\lambda \in \boldsymbol{K}$ が $A$ の**固有値**であるとは，

$$A\mathbf{u} = \lambda\mathbf{u}$$

をみたす $\mathbf{u} \in \boldsymbol{K}^N$, $\mathbf{u} \neq \mathbf{0}$ が存在することである．この関係式をみたす $\mathbf{0}$ でない $\mathbf{u}$ を $\lambda$ に関する $A$ の**固有ベクトル**という．

$A$ の固有値を求めるために，次の定理は非常に有用である．

**定理 13.6** $N$ 次正方行列 $A$ に対して，

$$\det\,(\lambda I_N - A) = 0 \iff \lambda \text{は} A \text{の固有値}.$$

**証明.** $\lambda$ は $A$ の固有値 $\iff$ ある $\mathbf{x} \neq \mathbf{0}$ が存在し，$A\mathbf{x} = \lambda\mathbf{x} \iff$ ある $\mathbf{x} \neq \mathbf{0}$ が存在し，$(A - \lambda I_N)\mathbf{x} = \mathbf{0} \iff A - \lambda I_N$ は正則でない $\iff |A - \lambda I_N| = 0$. ■

**定義 13.4** 第 7.4 節で定義したが，$N$ 次正方行列 $A$ に対して，

$$\Delta_A(z) = \det\,(zI_N - A) = (-1)^N \det\,(A - zI_N)$$

を $A$ の**固有多項式 (特性多項式)** という．また方程式

$$\Delta_A(z) = 0$$

を**固有方程式**(**特性方程式**) という．

$N$ 次正方行列 $A$ に対して $\Delta_A(z)$ は $z$ に関する $\boldsymbol{K}$ に属する数を係数にもつ $N$ 次多項式，すなわち

$$\Delta_A(z) = z^N + c_{N-1}z^{N-1} + \cdots + c_1 z + c_0 \ \ (c_0, \cdots, c_{N-1} \in \boldsymbol{K})$$

となっている．次の代数学の基本定理が知られている (本書ではこの定理の証明は省略する．詳細はたとえば [野口] を参照してほしい).

**定理 13.7**（**代数学の基本定理**）　$\boldsymbol{K}$ に属する数を係数にもつ $N$ 次多項式 $p(z)$ に対して，ある $\lambda_1, \cdots, \lambda_N \in \boldsymbol{C}$ と $c \in \boldsymbol{K}$ が存在し，すべての $z \in \boldsymbol{K}$ に対して

$$p(z) = c\,(z - \lambda_1) \cdots (z - \lambda_N)\,.$$

(注：$\boldsymbol{K} = \boldsymbol{R}$ の場合でも一般には $\lambda_1, \cdots, \lambda_N \in \boldsymbol{C}$ である．たとえば $z^2 + 1 = (z - \sqrt{-1})\,(z + \sqrt{-1})$ である．)

各 $\lambda_i$ は $p(\lambda_i) = 0$ をみたすので, $p(z) = 0$ の解である．一般にこれら $\lambda_1, \cdots, \lambda_N$ の中には同じものが含まれている場合もある．特に $\lambda_{i(1)} = \cdots = \lambda_{i(r)} = \lambda'$ であり，$\lambda_j \neq \lambda' \ (j \notin \{i(1), \cdots, i(r)\})$ であるとき，$\lambda'$ は $p(z) = 0$ の**重複度** $r$ の解という．定理 13.6 から固有多項式 $\Delta_A(z) = 0$ の解は $A$ の固有値であるが，特に重複度 $r$ の解であるとき，**重複度** $r$ **の固有値**という．

代数学の基本定理から次のことが成り立つ．

**系 13.8**　$N$ 次正方行列は，必ず重複度までこめて，$N$ 個の固有値 $\lambda_1, \cdots, \lambda_N \in \boldsymbol{K}$ をもつ．

固有値，固有ベクトルと行列の対角化には次のような密接な関係がある．

**定理 13.9**　$A$ を $N$ 次正方行列とする．次の (1), (2) が成り立つ．

(1)　$A$ が対角化できるとする．すなわち $\boldsymbol{K}^N$ のある基底 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ と $\lambda_1, \cdots, \lambda_N \in \boldsymbol{K}$ が存在し，

$$X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1}\, A\, X_{\mathcal{E}\leftarrow\mathcal{U}} = (\mathbf{u}_1, \cdots, \mathbf{u}_N)^{-1}\, A\, (\mathbf{u}_1, \cdots, \mathbf{u}_N)$$

$$= \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_N \end{pmatrix} \tag{13.6}$$

が成り立つとする．このとき，$\lambda_1, \cdots, \lambda_N$ は $A$ の重複度までいれた $N$ 個の固有値で，$\mathbf{u}_i$ は $\lambda_i$ の固有ベクトルである, $i = 1, \cdots, N$.

(2)　$\lambda_1, \cdots, \lambda_N$ を $A$ の重複度までいれた $N$ 個の固有値とする．もし $\lambda_j$ に関する固有ベクトル $\mathbf{u}_i$, $i = 1, \cdots, N$ を $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の基底になるようにとれるとする．このとき，$A$ は (13.6) のように対角化できる．

**証明.** (1)　$\boldsymbol{K}^N$ のある基底 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ をとって，

$$A\ (\mathbf{u}_1, \cdots, \mathbf{u}_N) = (\mathbf{u}_1, \cdots, \mathbf{u}_N) \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_N \end{pmatrix} \tag{13.7}$$

が成立している．このことから，

$$(A\mathbf{u}_1, \cdots A\mathbf{u}_N) = (\lambda_1 \mathbf{u}_1,\ \cdots,\ \lambda_N \mathbf{u}_N) \tag{13.8}$$

となることがわかる．したがって，$\lambda_j$ は $A$ の固有値であり，$\mathbf{u}_j$ は $\lambda_j$ に関する固有ベクトルである．$P = X_{\mathcal{E} \leftarrow \mathcal{U}}\ (= (\mathbf{u}_1, \cdots, \mathbf{u}_N))$ とおくと，特性多項式が

$$\Delta_A(z) = \Delta_{P^{-1}AP}(z) = \Delta_{\mathrm{diag}(\lambda_1, \cdots, \lambda_N)}(z) = (z - \lambda_1) \cdots (z - \lambda_N)$$

となり，$\lambda_1, \cdots, \lambda_N$ が $A$ の重複度までいれた $N$ 個の固有値であることがわかる．

(2)　$\mathbf{u}_j$ が $\lambda_j$ の固有ベクトルであることから，(13.7) が成り立ち，$X_{\mathcal{E} \leftarrow \mathcal{U}} = (\mathbf{u}_1, \cdots, \mathbf{u}_n)$ であるから，(13.6) が成立する．■

この定理から，次のことがわかる．

**行列$A$ が対角化可能であるとき，$A$ を対角化するような基底，あるいは正則行列を見出すには，$A$ の固有値と固有ベクトルを求めればよい．**

一般に与えられた行列の固有値と固有ベクトルを求めることを**固有値問題**という．

**例題 13.2**　$A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$ の固有値と固有ベクトルを求め，対角化せよ．

**答．** $A\mathbf{x} = \lambda\mathbf{x}$ が成り立つことと

$$(A - \lambda I_2)\,\mathbf{x} = \begin{pmatrix} 2-\lambda & 1 \\ 3 & 4-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{13.9}$$

とは同値であるから，(13.9) の非自明解を求めればよい．(13.9) が非自明解をもつための必要十分条件は

$$|A - \lambda I_2| = \left| \begin{pmatrix} 2-\lambda & 1 \\ 3 & 4-\lambda \end{pmatrix} \right| = (\lambda - 1)(\lambda - 5) = 0$$

となることである．したがって，$\lambda = 1, 5$ の場合に非自明解をもつ．ゆえに $1, 5$ が $A$ の固有値である．次に固有ベクトルを求める．$\lambda = 1$ の場合，

$$\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

の非自明解を求めればよい．行に関する基本変形で

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 3 & 3 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

となる．したがって，$x_2 = s$, $x_1 = -s$ が解，すなわち

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad s \neq 0$$

が 1 に関する固有ベクトルである．$\lambda = 5$ の場合，

$$\begin{pmatrix} -3 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

の非自明解を求めればよい．行に関する基本変形により，

$$\left(\begin{array}{cc|c} -3 & 1 & 0 \\ 3 & -1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} -3 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{array}\right)$$

となるから，$x_2 = s$, $x_1 = \dfrac{1}{3}s$，すなわち

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s \begin{pmatrix} \frac{1}{3} \\ 1 \end{pmatrix}, \quad s \neq 0$$

が 5 に関する固有ベクトルである．そこで

$$\mathbf{u}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

とする．これが基底になっていることを示す．行に関する基本変形により

$$(\mathbf{u}_1, \mathbf{u}_1) = \begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{-1} & 1 \\ 0 & \boxed{4} \end{pmatrix}$$

となり，ピボットが 2 個あるので，$\mathrm{rank}\,(\mathbf{u}_1, \mathbf{u}_1) = 2$，すなわち $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_1\}$ は基底になっている．よって

$$P = (\mathbf{u}_1, \mathbf{u}_1) = \begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix}$$

により $A$ は対角化できる．実際，

$$P^{-1}AP = \begin{pmatrix} -\frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

となる．■

対角化不可能な行列の例もあげておく．

**例 13.3** $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ は対角化可能でない．

**証明.** $\Delta_A(\lambda) = |\lambda I_2 - A| = \begin{vmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2$ であるから固有値は 1, 1 である．したがって，もし対角化できるならば，$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$ をみたす正則行列が存在することになる．これより $AP = PI_2 = P$ である．ゆえに $A = PP^{-1} = I_2$ でなければならないが，$A \neq I_2$ である．■

このように行列には対角化できるものと対角化できないものがある．以下では，どのような行列が対角化できるのか，行列が対角化できるかどうかを判定するための条件は何か，といったことについて述べていく．その前に準備として固有空間の概念を導入する．

**定義 13.5** $N$ 次正方行列 $A$ とその固有値 $\lambda$ に対して

$$V_\lambda = \left\{\mathbf{u} : \mathbf{u} \in \boldsymbol{K}^N,\ A\mathbf{u} = \lambda\mathbf{u}\right\}$$

を $\lambda$ に関する**固有空間**という．

**命題 13.10**　固有空間 $V_\lambda$ は $\boldsymbol{K}^N$ の部分空間である．

**証明.** $\mathbf{u}, \mathbf{v} \in V_\lambda$ に対して，

$$A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \lambda\mathbf{u} + \lambda\mathbf{v} = \lambda(\mathbf{u}+\mathbf{v})$$

であるから，$\mathbf{u}+\mathbf{v} \in V_\lambda$ である．また $\alpha \in \boldsymbol{K}$ に対して，

$$A(\alpha\mathbf{u}) = \alpha A\mathbf{u} = \alpha\lambda\mathbf{u} = \lambda(\alpha\mathbf{u})$$

より $\alpha\mathbf{u} \in V_\lambda$ が成り立つ．■

**定理 13.11**　$\lambda$ を $A$ の重複度 $r$ の固有値とする．このとき

$$\dim V_\lambda \leq r$$

である．

**証明.** $s = \dim V_\lambda$ とする．$V_\lambda$ の基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_s\}$ をとる．

このとき，$\mathbf{u}_{s+1}, \cdots, \mathbf{u}_N \in \boldsymbol{K}^N$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の基底となるように選ぶ．$P = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$ とすると

$$P^{-1}AP = P^{-1}(A\mathbf{u}_1, \cdots, A\mathbf{u}_N) = P^{-1}(\lambda\mathbf{u}_1, \cdots, \lambda\mathbf{u}_s, A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N) \tag{13.10}$$

であるが，$A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N \in \operatorname{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ であるから，

$$A\mathbf{u}_j = \sum_{k=1}^{N} c_{jk}\mathbf{u}_k$$

と表せる．ゆえに

$$\begin{aligned}
&(\lambda\mathbf{u}_1, \cdots, \lambda\mathbf{u}_s, A\mathbf{u}_{s+1}, \cdots, A\mathbf{u}_N) \\
&= \left(\lambda\mathbf{u}_1, \cdots, \lambda\mathbf{u}_s, \sum_{k=1}^{N} c_{s+1,k}\mathbf{u}_k, \cdots, \sum_{k=1}^{N} c_{N,k}\mathbf{u}_k\right)
\end{aligned}$$

$$= (\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{u}_{s+1}, \cdots, \mathbf{u}_N) \begin{pmatrix} \lambda & & & c_{s+1,1} & \cdots & c_{N,1} \\ & \ddots & & \vdots & & \vdots \\ & & \lambda & c_{s+1,s} & \cdots & c_{N,s} \\ & & & c_{s+1,s+1} & \cdots & c_{N,s+1} \\ & O & & \vdots & & \vdots \\ & & & c_{s+1,N} & \cdots & c_{N,N} \end{pmatrix}.$$

ゆえにブロック行列を用いて

$$\begin{aligned}(13.10) &= P^{-1}(\mathbf{u}_1, \cdots, \mathbf{u}_s, \mathbf{u}_{s+1}, \cdots, \mathbf{u}_N) \begin{pmatrix} \lambda I_s & B \\ O & C \end{pmatrix} \\ &= P^{-1}P \begin{pmatrix} \lambda I_s & B \\ O & C \end{pmatrix} = \begin{pmatrix} \lambda I_s & B \\ O & C \end{pmatrix}\end{aligned}$$

と表せる．したがって

$$\begin{aligned}\Delta_A(z) = \Delta_{P^{-1}AP}(z) &= \det \begin{pmatrix} (z-\lambda)I_s & -B \\ O & zI_{N-s} - C \end{pmatrix} \\ &= (z-\lambda)^s \det(zI_{N-s} - C)\end{aligned}$$

であるから，$\lambda$ の重複度は少なくとも $s$ 以上である．■

一般に $\dim V_\lambda = r$ になるとは限らない．しかし後に示すように，$A$ が対角化できるための必要十分条件は，すべての相異なる固有値 $\lambda$ に対して

$$\dim V_\lambda = \text{“}\lambda\text{の重複度”}$$

となることである (定理 13.15)．まず $\dim V_\lambda < r$ となる例をあげておこう．

**例 13.4** $A = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ $(c \neq 0)$ とする．$\Delta_A(\lambda) = (\lambda - 1)^2$ である．1 が $A$ の重複度 2 の固有値である．1 の固有空間を求める．

$$A - I_2 = \begin{pmatrix} 0 & -c \\ 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & \boxed{1} \\ 0 & 0 \end{pmatrix}$$

であるから，$x_2 = 0, x_1 = s$ が解である．したがって

$$V_1 = \left\{ s \begin{pmatrix} 1 \\ 0 \end{pmatrix} : s \in \boldsymbol{K} \right\}$$

であり，$\dim V_1 = 1$.

**定理 13. 12**　$\lambda_1, \cdots, \lambda_s$ を $A$ の相異なる固有値とする．$\mathbf{u}_i$ を $\lambda_i$ の一つの固有ベクトルとする $(i = 1, \cdots, s)$．このとき，$\mathbf{u}_1, \cdots, \mathbf{u}_s$ は線形独立である．

**証明.** $\mathbf{u}_1, \mathbf{u}_2$ は線形独立である．なぜなら，もし線形従属ならば $\mathbf{u}_1 = k\mathbf{u}_2$ と表せるが，

$$\lambda_1 \mathbf{u}_1 = A\mathbf{u}_1 = kA\mathbf{u}_2 = k\lambda_2 \mathbf{u}_2 = \lambda_2 \mathbf{u}_1$$

となり，$\lambda_1 \neq \lambda_2$ に矛盾する．もし $\mathbf{u}_1, \cdots, \mathbf{u}_s$ が線形従属であるとすると，少なくとも $\mathbf{u}_1, \mathbf{u}_2$ は線形独立であるから，$\mathbf{u}_1, \cdots, \mathbf{u}_t$ が線形独立であるが，$\mathbf{u}_1, \cdots, \mathbf{u}_{t+1}$ が線形従属となる $t$ が存在する．このとき，

$$\mathbf{u}_{t+1} = c_1 \mathbf{u}_1 + \cdots + c_t \mathbf{u}_t$$

と表せる．$c_1, \cdots, c_t$ のうち 0 でないものが必ず存在するので，その一つを $c_k$ とおく．両辺に $A$ を作用させれば

$$\lambda_{t+1} \mathbf{u}_{t+1} = A\mathbf{u}_{t+1} = \sum_{j=1}^{t} c_j A\mathbf{u}_j = \sum_{j=1}^{t} c_j \lambda_j \mathbf{u}_j.$$

一方，$\lambda_{t+1} \mathbf{u}_{t+1} = \sum_{j=1}^{t} \lambda_{t+1} c_j \mathbf{u}_j$ であるから

$$\mathbf{0} = \lambda_{t+1} \mathbf{u}_{t+1} - \lambda_{t+1} \mathbf{u}_{t+1} = \sum_{j=1}^{t} (\lambda_j c_j - \lambda_{t+1} c_j)\, \mathbf{u}_j.$$

ここで $\mathbf{u}_1, \cdots, \mathbf{u}_t$ は線形独立であるから，$\lambda_j c_j = \lambda_{t+1} c_j$ である．$c_k \neq 0$ であるから $\lambda_k = \lambda_{t+1}$ となり，相異なる固有値をとったことに反する．■

**系 13. 13**　$\lambda_1, \cdots, \lambda_s$ を $A$ の相異なる固有値とする．$i = 1, \cdots, s$ に対して，$\left\{\mathbf{x}_1^{(i)}, \cdots, \mathbf{x}_{N_i}^{(i)}\right\}$ を $V_{\lambda_i}$ の基底とする．このとき，

$$\mathbf{x}_1^{(1)},\ \cdots,\ \mathbf{x}_{N_1}^{(1)},\ \mathbf{x}_1^{(2)},\ \cdots,\ \mathbf{x}_{N_2}^{(2)},\ \cdots,\ \mathbf{x}_1^{(s)},\ \cdots,\ \mathbf{x}_{N_s}^{(s)}$$

は線形独立である．

**証明.** $c_1^{(1)} \mathbf{x}_1^{(1)} + \cdots + c_{N_s}^{(s)} \mathbf{x}_{N_s}^{(s)} = \mathbf{0}$ とする．このとき，$i = 1, \cdots, s$ に対して

$$\mathbf{y}_i = c_1^{(i)} \mathbf{x}_1^{(i)} + \cdots + c_{N_i}^{(i)} \mathbf{x}_{N_i}^{(i)}$$

とおくと，$\mathbf{y}_i \in V_{\lambda_i}$ であり，かつ $\sum_{i=1}^{s} \mathbf{y}_i = \mathbf{0}$ となる．定理 13.12 より $\mathbf{y}_i = 0\,(i =$

$1,\cdots,s)$ である．ゆえに各 $i$ に対して $\sum_{n=1}^{N_i} c_n^{(i)}\mathbf{x}_n^{(i)} = 0$ となり，$c_n^{(i)} = 0\,(i = 1,\cdots,s,\ n = 1,\cdots,N_i)$ が得られる．■

**系 13.14** $A$ を $N$ 次正方行列とする．もし $A$ の固有値がすべて相異なるならば，$A$ は対角化できる．

**証明．** $\lambda_1,\cdots,\lambda_N$ を $A$ の重複度もいれた $N$ 個の固有値とする．仮定よりこれらの固有値はすべて相異なる．$\lambda_i$ の固有ベクトル $\mathbf{u}_i$ をとる．定理 13.12 より $\mathbf{u}_1,\cdots,\mathbf{u}_N$ は線形独立であるから，$\{\mathbf{u}_1,\cdots,\mathbf{u}_N\}$ は $\boldsymbol{K}^N$ の基底になっている．ゆえに定理 13.9 より $A$ は対角化可能である．■

固有値がすべて相異ならなくとも対角化できる場合がある．対角化できるための必要十分条件をあげる．

**定理 13.15** $A$ を $N$ 次正方行列とし，$\lambda_1,\cdots,\lambda_s$ を $A$ の相異なる固有値とする．次の (1), (2), (3) は同値である．

(1) $A$ は対角化可能である．

(2) $\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_s} = N$.

(3) $i = 1,\cdots,s$ に対して，$\dim V_{\lambda_i}$ は $\lambda_i$ の重複度と一致している．

**証明．** (1) ⇒ (3)　$\lambda_i$ の重複度を $r_i$ とする $(i = 1,\cdots,s)$．定理 13.9 より $\lambda_i$ に関する固有ベクトル $\mathbf{u}_1^{(i)},\cdots,\mathbf{u}_{r_i}^{(i)}$ で，

$$\mathbf{u}_1^{(1)},\ \cdots,\ \mathbf{u}_{r_1}^{(1)},\ \cdots,\ \mathbf{u}_1^{(s)},\ \cdots,\ \mathbf{u}_{r_s}^{(s)}$$

が $\boldsymbol{K}^N$ の基底になるものをとることができる．$\mathbf{u}_1^{(i)},\cdots,\mathbf{u}_{r_i}^{(i)} \in V_{\lambda_i}$ であり，線形独立である．定理 13.11 より，$\dim V_\lambda \leq r_i$ であるから，$\left\{\mathbf{u}_1^{(i)},\cdots,\mathbf{u}_{r_i}^{(i)}\right\}$ は $V_{\lambda_i}$ の基底であり，

$$\dim V_{\lambda_i} = r_i \tag{13.11}$$

である．ゆえに (3) が成り立つ．

(3) ⇒ (2) は明らかである．

(2) ⇒ (1)　(2) と定理 13.11 より $\dim V_{\lambda_i} = r_i,\ i = 1,\cdots,s$ であることがわかる．$V_{\lambda_i}$ の基底 $\left\{\mathbf{u}_1^{(i)},\cdots,\mathbf{u}_{r_i}^{(i)}\right\}$ をとると，$\mathbf{u}_1^{(1)},\cdots,\mathbf{u}_{r_1}^{(1)},\cdots,\mathbf{u}_1^{(s)},\cdots,\mathbf{u}_{r_s}^{(s)}$ は系 13.13 より，線形独立である．また (2) よりこれらは $N$ 個のベクトルであるから，$\boldsymbol{K}^N$ の基底になっている．したがって，

$$P = \left(\mathbf{u}_1^{(1)}, \cdots, \mathbf{u}_{r_1}^{(1)}, \cdots, \mathbf{u}_1^{(s)}, \cdots, \mathbf{u}_{r_s}^{(s)}\right)$$

は正則行列で，定理 13.9 の証明と同様にして，$P^{-1}AP$ が対角行列であることがわかる．■

**例題 13.3**　$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ が対角化できるかどうかを調べ，対角化できる場合は対角化せよ．

**答．** $\Delta_A(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda + 1)$ であるから，固有値は $-1, 1, 2$ である．固有値はすべて相異なるので対角化できる．$-1$ の固有空間を求める．行に関する基本変形により

$$A + I_3 = \begin{pmatrix} 2 & 0 & 2 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

となる．したがって，$x_3 = s, x_2 = 0, x_1 = -s$ が $(A + I_3)\mathbf{x} = \mathbf{0}$ の解である．ゆえに

$$V_{-1} = \left\{ s \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} : s \in \boldsymbol{K} \right\}.$$

次に 1 に関する固有空間を求める．行に関する基本変形により，

$$A - I_3 = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 0 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix}.$$

ゆえに $x_2 = s, x_3 = 0, x_1 = 0$ が $(A - I_3)\mathbf{x} = 0$ の解である．ゆえに

$$V_1 = \left\{ s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : s \in \boldsymbol{K} \right\}$$

が固有空間である．2 に関する固有空間を求める．

$$A - 2I_3 = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & -2 \\ 0 & \boxed{1} & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

したがって，$x_3 = s,\ x_2 = 3s,\ x_1 = 2s$ が $(A - 2I_3)\,\mathbf{x} = \mathbf{0}$ の解である．ゆえに

$$V_2 = \left\{ s \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} : s \in \boldsymbol{K} \right\}$$

である．いま

$$\mathbf{u}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

はそれぞれ $-1, 1, 2$ の固有ベクトルである．$P = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ とし，これの逆行列を求める．

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{3} & 0 & \frac{1}{3} \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{3} & 0 & \frac{1}{3} \end{array}\right).$$

ゆえに

$$P^{-1}AP = \begin{pmatrix} -\frac{1}{3} & 0 & \frac{2}{3} \\ -1 & 1 & -1 \\ \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 2 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ について調べる．例 13.4 より，1 が重複度 2 の固有値であり，一方，$\dim V_1 = 1$ であるから，定理 13.15 より $B$ は対角化できない． ■

**問題 13.2** 次の行列が対角化可能かどうかを調べ，対角化可能の場合には対角化せよ．

$$A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

## 13.3　正規直交基底による対角化

$N$ 次正方行列 $A$ が対角化できるとは，$\boldsymbol{K}^N$ のある基底 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ をとって，

$$X_{\mathcal{E}\leftarrow\mathcal{U}}^{-1}\, A\, X_{\mathcal{E}\leftarrow\mathcal{U}} = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)$$

とできることであったが，ここでは特に正規直交基底 $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ によって対角化できる行列について述べる．

**定義 13.6**　$T = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$ を $N$ 次正方行列とする．$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の正規直交基底であるとき，$T$ を**ユニタリー行列**という．

**定義 13.7**　特に $\boldsymbol{K}^N = \boldsymbol{R}^N$ の場合は，ユニタリー行列のことを**直交行列**という．したがって，正確には，$\boldsymbol{K}^N = \boldsymbol{C}^N$ の場合にユニタリー行列という．ここでは記述を簡単にするため，ユニタリー行列という言葉で統一している．

ユニタリー行列は次のように特徴付けられる．

**定理 13.16**　$T$ を $N$ 次正方行列とする．このとき，次の (1)～(4) は同値である．

(1)　$T$ はユニタリー行列である．

(2)　$T^*T = I_N$.

(3)　$T^*T = TT^* = I_N$.

(4)　任意の $\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\langle T\mathbf{x}, T\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$$

が成り立つ．特に

$$\|T\mathbf{x}\| = \|\mathbf{x}\|$$

である．

**証明.** (1) $\Leftrightarrow$ (2) を示す．

$$T^*T = \begin{pmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_n^* \end{pmatrix} (\mathbf{u}_1, \cdots, \mathbf{u}_n) = \begin{pmatrix} \mathbf{u}_1^*\mathbf{u}_1 & \cdots & \mathbf{u}_1^*\mathbf{u}_n \\ \vdots & \ddots & \vdots \\ \mathbf{u}_n^*\mathbf{u}_1 & \cdots & \mathbf{u}_n^*\mathbf{u}_n \end{pmatrix}$$

$$= \begin{pmatrix} \overline{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} & \cdots & \overline{\langle \mathbf{u}_1, \mathbf{u}_n \rangle} \\ \vdots & \ddots & \vdots \\ \overline{\langle \mathbf{u}_n, \mathbf{u}_1 \rangle} & \cdots & \overline{\langle \mathbf{u}_n, \mathbf{u}_n \rangle} \end{pmatrix}$$

であるから，$\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ が正規直交基底であることと $T^*T = I_N$ は同値である．

(2) $\Rightarrow$ (3) を示す．(2) より，$T$ は正則であり，$T^* = T^{-1}$ である．ゆえに $I_N = TT^{-1} = TT^*$ でもある．(3) $\Rightarrow$ (2) は明らかである．

$T$ がユニタリー行列ならば，

$$\overline{\langle T\mathbf{x}, T\mathbf{y} \rangle} = (T\mathbf{x})^* T\mathbf{y} = \mathbf{x}^*T^*T\mathbf{y} = \mathbf{x}^*\mathbf{y} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$$

であるから，(1) $\Rightarrow$ (4) が成り立つ．逆に，(4) が成り立つならば，

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} - T^*T\mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, T^*T\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle - \langle T\mathbf{x}, T\mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle = 0 \end{aligned}$$

であるから，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$\langle \mathbf{x}, \mathbf{y} - T^*T\mathbf{y} \rangle = 0.$$

これより $\mathbf{y} = T^*T\mathbf{y}$ を得る．したがって (4) $\Rightarrow$ (2) が成り立つ． ■

**系 13. 17** $T$ がユニタリー行列ならば，$T^* = T^{-1}$ である．

**系 13. 18** $U_1, U_2$ を $N$ 次ユニタリー行列とする．このとき $U_1U_2$ は $N$ 次ユニタリー行列である．

**証明.** $(U_1U_2)^*(U_1U_2) = U_2^*U_1^*U_1U_2 = I_N$. ■

**系 13. 19** $T$ がユニタリー行列ならば $|T| = \pm 1$ である．

**証明.** $\det T^* = \overline{\det T}$ であるから $1 = \det T^* \det T = |\det T|^2$. ■

**注意 13. 1** $A$ を 2 次の直交行列とする．このときある $\theta \in [0, 2\pi)$ によって

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{または} \quad A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

と表せることを示す．実際 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とすると，$a^2+c^2=b^2+d^2=1$ である．したがって適当な $\theta\in[0,2\pi)$ をとって $a=\cos\theta$, $c=\sin\theta$ と表せる．

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

であるが，一方

$$A^{-1}=A^T=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

でもある．したがって $|A|=1$ の場合，

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

であるから，$d=a$, $b=-c$ が成り立つ．ゆえに

$$A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{13.12}$$

と表せる．また $|A|=-1$ ならば

$$A=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{13.13}$$

である．なお，この二つの行列の間には

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

の関係式がある．(13.12) の場合，$A\mathbf{x}$ は $\mathbf{x}$ を $\mathbf{0}$ を中心に角度 $\theta$ だけ回転させたものである．

$$\mathbf{x}'=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\mathbf{x}=\begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

であるから，$\mathbf{x}'$ は $\mathbf{x}$ の $x_1$ 軸に対称な点を表している．したがって (13.13) の場合，$A\mathbf{x}$ は $\mathbf{x}'$ をさらに $\mathbf{0}$ を中心に角度 $\theta$ だけ回転させたものである．

ここで考える課題は次のものである．

**どのような $N$ 次正方行列 $A$ に対して，ユニタリー行列 $U$ で**

$$U^*AU=\mathrm{diag}\,(\lambda_1,\cdots,\lambda_N)$$

**となるものが存在するか？**

まず，ユニタリー行列 $U$ で対角化できたとする．$D = \mathrm{diag}(\lambda_1, \cdots, \lambda_N)$ とすると，$DD^* = D^*D$ であるが，$U$ がユニタリー行列であることより，

$$DD^* = (U^*AU)(U^*A^*U) = U^*AA^*U,$$
$$D^*D = (U^*A^*U)(U^*AU) = U^*A^*AU.$$

ゆえに

$$AA^* = A^*A$$

の関係式がみたされていることが必要である．じつは $A$ がこの関係式をみたしていれば，$A$ はユニタリー行列を用いて対角化することができる．本節ではこのことを証明する．

**定義 13.8**　$N$ 次正方行列 $A$ が $AA^* = A^*A$ をみたしているとき，$A$ を**正規行列**という．

正規行列の例をあげておく．

**例 13.5**　ユニタリー行列 $U$ は正規行列である．実際

$$UU^* = I_N = U^*U.$$

**例 13.6**　$N$ 次正方行列 $A$ が $A = A^*$ または $A = -A^*$ をみたすとき，$A$ は正規行列である．

特に例 13.6 の性質をもつ行列は応用上しばしば現れる行列で，次のように名前が付けられている．

**定義 13.9**　$N$ 次正方行列 $A$ が $A = A^*$ をみたすとき，**エルミート行列**という．また $A = -A^*$ をみたすとき**歪エルミート行列**という．なお特に $\boldsymbol{K} = \boldsymbol{R}$ の場合，エルミート行列は $A^T = A$ をみたす行列で，これを**実対称行列**という．またこの場合，歪エルミート行列は $A^T = -A$ をみたす行列で**実歪対称行列**と呼ばれる．

エルミート行列については次のことが成り立つ．

**定理 13.20**　$A$ を $N$ 次のエルミート行列とする．このとき，$A$ の固有値はすべて実数である．

**証明.** $\lambda$ を $A$ の固有値とし，$\mathbf{u}$ を $\lambda$ の長さ 1 の固有ベクトル，すなわち $\lambda$ の固有ベクトルで $\|\mathbf{u}\| = 1$ をみたすものとする．たとえば $\lambda$ の固有ベクトル $\mathbf{v}$ に対して $\mathbf{u} = \|\mathbf{v}\|^{-1}\mathbf{v}$ が長さ 1 の $\lambda$ の固有ベクトルになっている．このような $\mathbf{u}$ に対して

$$\begin{aligned}\lambda &= \langle \lambda\mathbf{u}, \mathbf{u}\rangle = \langle A\mathbf{u}, \mathbf{u}\rangle = \langle \mathbf{u}, A^*\mathbf{u}\rangle \\ &= \langle \mathbf{u}, A\mathbf{u}\rangle = \langle \mathbf{u}, \lambda\mathbf{u}\rangle = \overline{\lambda}.\end{aligned}$$

したがって，$\lambda$ は実数値である．■

エルミート行列の固有ベクトルは次の著しい性質を有している．

**定理 13.21**　$A$ をエルミート行列とする．$\lambda, \mu$ を $A$ の相異なる固有値とする．任意の $\mathbf{x} \in V_\lambda$ と $\mathbf{y} \in V_\mu$ は直交している．すなわち $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ が成立している．

**証明.** 前定理より $\lambda, \mu$ は実数である．ゆえに

$$\begin{aligned}(\lambda - \mu)\langle \mathbf{x}, \mathbf{y}\rangle &= \langle \lambda\mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{x}, \mu\mathbf{y}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{x}, A\mathbf{y}\rangle \\ &= \langle A\mathbf{x}, \mathbf{y}\rangle - \langle A\mathbf{x}, \mathbf{y}\rangle = 0\end{aligned}$$

である．よって $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ が得られる．■

はじめに，一般の正方行列がユニタリー行列により上三角行列にできることを示す．

**定理 13.22（シューアの定理）**　$A$ を $N$ 次正方行列とする．$\lambda_1, \cdots, \lambda_N$ を $A$ の重複度までいれた $N$ 個の固有値とする．このときユニタリー行列 $U$ をとって，

$$U^*AU = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ O & & \lambda_N \end{pmatrix}$$

のように上三角行列に変形できる．

**証明.** $N$ に関する帰納法で示す．$N = 1$ のときは明らか．$N - 1$ のとき成り立つとして $N$ の場合を示す．

$\lambda_1$ の長さ 1 の固有ベクトルを一つとり，$\mathbf{u}_1$ とする．このとき，適当に $\mathbf{u}_2, \cdots, \mathbf{u}_N$ を選んで，$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が $\boldsymbol{K}^N$ の正規直交基底になるようにできる．$U_1 = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$ とすると，$A\mathbf{u}_2, \cdots, A\mathbf{u}_N \in \mathrm{span}\,\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ であるから，

$$A\mathbf{u}_j = \sum_{k=1}^{N} \alpha_{kj}\mathbf{u}_k$$

と表せる．したがって

$$\begin{aligned}AU_1 &= (\lambda_1\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_N) = \left(\lambda_1\mathbf{u}_1, \sum_{k=1}^{N}\alpha_{k2}\mathbf{u}_k, \cdots, \sum_{k=1}^{N}\alpha_{kN}\mathbf{u}_k\right)\\ &= (\mathbf{u}_1, \cdots, \mathbf{u}_N)\begin{pmatrix}\lambda_1 & C\\ O & B\end{pmatrix}\end{aligned}$$

とあるブロック行列 $C, B$ を用いて表せる．ここで，$U_1^* = U_1^{-1}$ であるから

$$U_1^*AU_1 = \begin{pmatrix}\lambda_1 & C\\ O & B\end{pmatrix}$$

が成り立つ．また

$$\begin{aligned}(z-\lambda_1)\cdots(z-\lambda_N) &= \Delta_A(z) = \Delta_{U_1^*AU_1}(z)\\ &= \det\begin{pmatrix}z-\lambda_1 & -C\\ O & zI_{N-1}-B\end{pmatrix} = (z-\lambda_1)\det(zI_{N-1}-B)\end{aligned}$$

となっている．したがって $\det(zI_{N-1}-B) = (z-\lambda_2)\cdots(z-\lambda_N)$ が得られる．ゆえに $B$ の重複度までいれた固有値は $\lambda_2, \cdots, \lambda_N$ である．帰納法の仮定から，$N-1$ 次のあるユニタリー行列 $Q$ が存在し，

$$Q^*BQ = \begin{pmatrix}\lambda_2 & & *\\ & \ddots & \\ O & & \lambda_N\end{pmatrix}$$

とできる．そこで

$$U_2 = \begin{pmatrix}1 & O\\ O & Q\end{pmatrix} \in \boldsymbol{K}^{N\times N}$$

とおくと，

$$U_2^*U_2 = \begin{pmatrix}1 & O\\ O & Q^*Q\end{pmatrix} = I_N$$

より，$U_2$ はユニタリー行列である．したがって $U = U_1U_2$ とすると $U$ もユニタリー行列であり，

$$U^*AU = U_2^*(U_1^*AU_1)U_2 = \begin{pmatrix}1 & O\\ O & Q^*\end{pmatrix}\begin{pmatrix}\lambda_1 & C\\ O & B\end{pmatrix}\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}$$

$$
= \begin{pmatrix} \lambda_1 & C \\ O & Q^*B \end{pmatrix} \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & CQ \\ O & Q^*BQ \end{pmatrix}
$$

$$
= \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ O & & \lambda_N \end{pmatrix}.
$$

よって $N$ のときも成り立つ. ■

特に $A$ が正規行列の場合は次のことが成り立つ.

**定理 13.23**　$A$ を $N$ 次正方行列とする. $\lambda_1, \cdots, \lambda_N$ を $A$ の重複度までいれた $N$ 個の固有値とする. 次の (1), (2) は同値である.

(1)　$A$ は正規行列である.

(2)　あるユニタリー行列 $U$ をとって,

$$
U^*AU = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)
$$

のように対角化できる.

**証明.** (2) ⇒ (1) はすでに証明したので, (1) ⇒ (2) を示す. 定理 13.22 より, あるユニタリー行列 $U$ により

$$
U^*AU = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1N} \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & b_{N-1,N} \\ O & & & \lambda_N \end{pmatrix}
$$

とできる. $T = U^*AU$ とおく. このとき,

$$
T^* = U^*A^*U = \begin{pmatrix} \overline{\lambda_1} & & & O \\ \overline{b_{12}} & \overline{\lambda_2} & & \\ \vdots & \ddots & \ddots & \\ \overline{b_{1N}} & \cdots & \overline{b_{N-1,N}} & \overline{\lambda_N} \end{pmatrix}
$$

である. $A$ は正規行列であるから,

$$
\begin{aligned}
TT^* &= U^*AUU^*A^*U = U^*AA^*U = U^*A^*AU \\
&= U^*A^*UU^*AU = T^*T
\end{aligned}
$$

である．いま $TT^*$ の $(1,1)$ 成分は

$$|\lambda_1|^2+|b_{12}|^2+\cdots+|b_{1N}|^2$$

である．一方 $T^*T$ の $(1,1)$ 成分は $|\lambda_1|^2$ であるから，$b_{12}=\cdots=b_{1N}=0$ となっている．次に $TT^*$ の $(2,2)$ 成分は,

$$|\lambda_2|^2+|b_{23}|^2+\cdots+|b_{2N}|^2$$

であり，一方 $T^*T$ の $(2,2)$ 成分は $|b_{12}|^2+|\lambda_2|^2=|\lambda_2|^2$ であるから，$b_{23}=\cdots=b_{2N}=0$ が得られる．この議論を繰り返せば，$b_{ij}=0$ となり，$U^*AU=\mathrm{diag}\,(\lambda_1,\cdots,\lambda_N)$ が導かれる．■

**例題 13.4**　$a$ を 0 でない複素数とする．$A=\begin{pmatrix}-a & a\\ a & -a\end{pmatrix}$ が正規行列であることを示せ．また $A$ をユニタリー行列により対角化せよ．

**答．** $A^*=\begin{pmatrix}-\overline{a} & \overline{a}\\ \overline{a} & -\overline{a}\end{pmatrix}$ であるから,

$$AA^*=\begin{pmatrix}2\,|a|^2 & -2\,|a|^2\\ -2\,|a|^2 & 2\,|a|^2\end{pmatrix}=A^*A$$

より，正規行列である．固有多項式は

$$\Delta_A(z)=\begin{vmatrix}z+a & -a\\ -a & z+a\end{vmatrix}=z(z+2a)$$

と因数分解できるので，$z=0,\ -2a$ が固有値である．0 の固有空間 $V_0$ を求める．行に関する基本変形により

$$\begin{pmatrix}a & -a\\ -a & a\end{pmatrix}\longrightarrow\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\longrightarrow\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}$$

とできるので，$x_2=s,\ x_1=x_2=s$ が $A\mathbf{x}=\mathbf{0}$ の解であり，したがって,

$$V_0=\left\{s\begin{pmatrix}1\\ 1\end{pmatrix}:s\in\boldsymbol{C}\right\}$$

である．次に $-2a$ の固有空間 $V_{-2a}$ を求める．行に関する基本変形により

$$\begin{pmatrix}a & a\\ a & a\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$$

とできるので，$x_2=s,\ x_1=-s$ は $(A+2aI)\mathbf{x}=\mathbf{0}$ の解である．したがって,

$$V_{-2a} = \left\{ s \begin{pmatrix} -1 \\ 1 \end{pmatrix} : s \in \boldsymbol{C} \right\}$$

が固有空間である．$V_0$ および $V_{-2a}$ の長さ 1 のベクトルとして

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

を考える．これらは直交しているので，$U = (\mathbf{u}_1, \mathbf{u}_2)$ はユニタリー行列である．これにより $A$ は対角化できる．実際

$$U^*AU = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} -a & a \\ a & -a \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & -2a \end{pmatrix}.$$

■

**例題 13.5**　次の行列を直交行列により対角化せよ．

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

**答．** 固有値を求める．

$$\Delta_A(t) = \begin{vmatrix} t-2 & -1 & -1 \\ -1 & t-2 & -1 \\ -1 & -1 & t-2 \end{vmatrix} = \begin{vmatrix} t-4 & -1 & -1 \\ t-4 & t-2 & -1 \\ t-4 & -1 & t-2 \end{vmatrix}$$

$$= (t-4) \begin{vmatrix} 1 & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{vmatrix} = (t-4) \begin{vmatrix} 1 & 0 & 0 \\ 1 & t-1 & 0 \\ 1 & 0 & t-1 \end{vmatrix}$$

$$= (t-1)^2(t-4).$$

したがって固有値は $1, 1, 4$ である．1 の固有空間を求める．

$$A - I = \begin{pmatrix} 2-1 & 1 & 1 \\ 1 & 2-1 & 1 \\ 1 & 1 & 2-1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

であるから，$x_3 = s$, $x_2 = t$, $x_1 = -s - t$ が $(A - I)\mathbf{x} = \mathbf{0}$ の解である．したがって，

$$V_1 = \left\{ s \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} : s, t \in \boldsymbol{R} \right\}$$

が 1 の固有空間である．$\mathbf{p}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{p}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ は $V_1$ の基底であるが，これをシュミットの直交化により $V_1$ の正規直交基底にする．まず

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{p}_1\|}\mathbf{p}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

とする．次に

$$\mathbf{v}_2 = \mathbf{p}_2 - \langle \mathbf{p}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

ゆえに

$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

とすると，$\{\mathbf{u}_1, \mathbf{u}_2\}$ は $V_1$ の正規直交基底である．次に 4 の固有空間を求める．

$$A - 4I = \begin{pmatrix} 2-4 & 1 & 1 \\ 1 & 2-4 & 1 \\ 1 & 1 & 2-4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

であるから，$x_3 = s,\ x_2 = s,\ x_1 = s$ が $(A - 4I)\mathbf{x} = \mathbf{0}$ の解である．ゆえに

$$V_4 = \left\{ s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} : \boldsymbol{R} \right\}$$

である．ゆえに

$$\mathbf{u}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

が $V_2$ の長さ 1 のベクトルである．このとき

$$U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

は直交行列で，

$$\begin{aligned} &U^t AU \\ &= \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \end{aligned}$$

■

**定義 13.10**　$A$ が $N$ 次の正規行列であるとする．このとき，ある $N$ 次のユニタリー行列 $U = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$ が存在し，$U^* AU$ を対角行列 $\mathrm{diag}(\lambda_1, \cdots, \lambda_N)$ と

することができる．ただしここで $\lambda_1, \cdots, \lambda_N$ は $A$ の固有値である．これより

$$A = U \,\mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)\; U^* = \sum_{j=1}^{N} \lambda_j\, \mathbf{u}_j\, \mathbf{u}_j^*$$

と表せる．

$$A = \sum_{j=1}^{N} \lambda_j\, \mathbf{u}_j\, \mathbf{u}_j^*$$

を $A$ の**スペクトル分解**という．

**問題 13. 3**　次の行列を直交行列で対角化せよ．

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

## 13. 4　エルミート形式とクーラン-フィッシャーの定理

**定義 13. 11**　$A = (a_{ij}) \in \boldsymbol{K}^{N \times N}$ をエルミート行列とする．このとき $\mathbf{x} = (x_j) \in \boldsymbol{K}^N$ に対して，

$$Q_A(\mathbf{x}) = \sum_{j,k=1}^{N} a_{jk} \overline{x_j} x_k$$

を $A$ によって定まる**エルミート形式**という．特に $A = (a_{ij}) \in \boldsymbol{R}^{N \times N}$ が実対称行列のときは，$\mathbf{x} = (x_j) \in \boldsymbol{R}^N$ に対して，

$$Q_A(\mathbf{x}) = \sum_{j,k=1}^{N} a_{jk} x_j x_k$$

を $A$ によって定まる **2 次形式**という．

エルミート形式はさまざまな分野で現れる．本書では後でカーブ・フィッティングの問題を論ずるときに用いる．本節ではエルミート形式のもつ性質と固有値との関連について述べる．

エルミート形式は内積を使って次のように表すことができる．

**定理 13. 24**　$\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$Q_A(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^* A\mathbf{x}.$$

**証明.** $A\mathbf{x} = \left(\sum_{k=1}^{N} a_{jk}x_k\right)$ であるから

$$\langle A\mathbf{x}, \mathbf{x}\rangle = \sum_{j=1}^{N}\left(\sum_{k=1}^{N} a_{jk}x_k\right)\overline{x_j} = Q_A(\mathbf{x}).$$

また

$$\langle A\mathbf{x}, \mathbf{x}\rangle = \mathbf{x}^* A\mathbf{x}.$$

■

**系 13.25**　$A \in \boldsymbol{K}^{N\times N}$ をエルミート行列とする．このとき任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して $Q_A(\mathbf{x})$ は実数である．

**証明.** 定理 13.24 より

$$\overline{Q_A(\mathbf{x})} = \overline{\langle A\mathbf{x}, \mathbf{x}\rangle} = \langle \mathbf{x}, A\mathbf{x}\rangle = \langle A^*\mathbf{x}, \mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{x}\rangle = Q_A(\mathbf{x})$$

となっている． ■

$N$ 次エルミート行列 $A$ の重複度も込めた $N$ 個の固有値はすべて実数である (定理 13.20)．それを大きさによって

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N \tag{13.14}$$

と並べる．

**定理 13.26**　$A$ を $N$ 次エルミート行列で，(13.14) によりその固有値を並べる．$B = \left\{\mathbf{x} \in \boldsymbol{K}^N : \|\mathbf{x}\| = 1\right\}$ とする．$\mathbf{x}_{\min}$ を $\lambda_1$ の長さ 1 の固有ベクトル，また $\mathbf{x}_{\max}$ を $\lambda_N$ の長さ 1 の固有ベクトルとする．このとき

$$\lambda_1 = \min_{\mathbf{x}\in B} Q_A(\mathbf{x}) = Q_A(\mathbf{x}_{\min}),$$

$$\lambda_N = \max_{\mathbf{x}\in B} Q_A(\mathbf{x}) = Q_A(\mathbf{x}_{\max}).$$

**証明.** $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_N)$ とする．あるユニタリー行列 $U$ が存在し，$U^*AU = D$ とできる．任意の $\mathbf{x} \in B$ に対して，$\mathbf{y} = U^*\mathbf{x}$ とおくと $\mathbf{x} = U\mathbf{y}$ より

$$\begin{aligned}\langle A\mathbf{x}, \mathbf{x}\rangle &= \langle U^*A\mathbf{x}, U^*\mathbf{x}\rangle = \langle U^*AU\mathbf{y}, U^*U\mathbf{y}\rangle \\ &= \langle D\mathbf{y}, \mathbf{y}\rangle.\end{aligned}$$

また $\|\mathbf{y}\| = \|\mathbf{x}\|$ であるから，

$$\langle A\mathbf{x}, \mathbf{x}\rangle \geq \min_{\mathbf{z}\in B} \langle D\mathbf{z}, \mathbf{z}\rangle = \min_{\mathbf{z}\in B} \sum_{j=1}^{N} \lambda_j |z_j|^2 \geq \lambda_1$$

が得られる．また

$$\langle A\mathbf{x}_{\min}, \mathbf{x}_{\min}\rangle = \langle \lambda_1 \mathbf{x}_{\min}, \mathbf{x}_{\min}\rangle = \lambda_1 \|\mathbf{x}_{\min}\|^2 = \lambda_1$$

である．ゆえに定理の前半が証明された．後半も同様に示すことができる． ■

$\lambda_1$ と $\lambda_N$ 以外の固有値についても同様の結果がいえる．

**定理 13.27（クーラン-フィッシャーの定理）** $A$ を $N$ 次エルミート行列で，(13.14) によりその固有値を並べる．$B = \left\{\mathbf{x} \in \boldsymbol{K}^N : \|\mathbf{x}\| = 1\right\}$ とする．$\mathcal{V}_k$ により，$\boldsymbol{K}^N$ の $k$ 次元部分空間全体のなす集合を表す．このとき次のことが成り立つ．

$$\lambda_j = \min_{V\in\mathcal{V}_j} \max_{\mathbf{x}\in B\cap V} Q_A(\mathbf{x}) = \max_{V\in\mathcal{V}_{N-j+1}} \min_{\mathbf{x}\in B\cap V} Q_A(\mathbf{x}) \tag{13.15}$$

**証明.** 行列 $A \in \boldsymbol{K}^{N\times N}$ と集合 $V \subset \boldsymbol{K}^N$ に対して，

$$A(V) = \{A\mathbf{x} : \mathbf{x} \in V\}$$

と表す．$U, D$ は定理 13.26 の証明で用いたものとする．まず (13.15) の左側の等式を証明する．$\mathbf{y} = U^*\mathbf{x}$ とおくと

$$\langle A\mathbf{x}, \mathbf{x}\rangle = \langle D\mathbf{y}, \mathbf{y}\rangle$$

であり，$V \subset \boldsymbol{K}^N$ に対して，$U^*(B \cap V) = B \cap U^*(V)$ であるから

$$\max_{\mathbf{x}\in B\cap V} \langle A\mathbf{x}, \mathbf{x}\rangle = \max_{\mathbf{y}\in B\cap U^*(V)} \langle D\mathbf{y}, \mathbf{y}\rangle$$

が成り立つ．また任意の $V \in \mathcal{V}_j$ に対して，$X = U^*(V)$ とすると，$X \in \mathcal{V}_j$ であり，$U(X) = V$ をみたしているので，

$$\inf_{V\in\mathcal{V}_j} \max_{\mathbf{x}\in B\cap V} \langle A\mathbf{x}, \mathbf{x}\rangle = \inf_{V\in\mathcal{V}_j} \max_{\mathbf{x}\in B\cap V} \langle D\mathbf{x}, \mathbf{x}\rangle$$

が成り立つ．以下では $\inf_{V\in\mathcal{V}_j} \max_{\mathbf{x}\in B\cap V} \langle D\mathbf{x}, \mathbf{x}\rangle = \lambda_j$ を証明する．$V \in \mathcal{V}_j$ を任意にとる．標準基底 $\mathbf{e}_1, \cdots, \mathbf{e}_N$ をとり，$W = \operatorname{span}\{\mathbf{e}_N, \mathbf{e}_{N-1}, \cdots, \mathbf{e}_j\}$ とする．このとき，$V \cap W \neq \{\mathbf{0}\}$ である．なぜなら，もし $V \cap W = \{\mathbf{0}\}$ とすると，

$$\begin{aligned} N &\geq \dim(V \dotplus W) = \dim V + \dim W = j + N - j + 1 \\ &= N + 1 \end{aligned}$$

となってしまうからである．したがって，$\mathbf{x} \in B \cap V \cap W$ が存在する．このよう

な任意の $\mathbf{x}$ は $\mathbf{x} = \sum_{k=j}^{N} x_k \mathbf{e}_k$, $\sum_{k=j}^{N} |x_k|^2 = 1$ と表される．ゆえに

$$\langle D\mathbf{x}, \mathbf{x} \rangle = \sum_{k=j}^{N} \lambda_k |x_k|^2 \geq \lambda_j$$

が成り立つ．これより

$$\max_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle \geq \max_{\mathbf{x} \in B \cap V \cap W} \langle D\mathbf{x}, \mathbf{x} \rangle \geq \lambda_j$$

となっている．したがって，$\inf_{V \in \mathcal{V}_j} \max_{\mathbf{x} \in B \cap V} (D\mathbf{x}, \mathbf{x}) \geq \lambda_j$ である．いま特に $V_0 = \operatorname{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\}$ とすると, 任意の $\mathbf{x} \in B \cap V_0$ に対して, $\mathbf{x} = \sum_{k=1}^{j} x_k \mathbf{e}_k$, $\sum_{k=1}^{j} |x_k|^2 = 1$ と表せるので,

$$\langle D\mathbf{x}, \mathbf{x} \rangle = \sum_{k=1}^{j} \lambda_k |x_k|^2 \leq \lambda_j$$

が得られる．ゆえに

$$\lambda_j \leq \inf_{V \in \mathcal{V}_j} \max_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle \leq \max_{\mathbf{x} \in B \cap V_0} \langle D\mathbf{x}, \mathbf{x} \rangle \leq \lambda_j$$

も成り立っている．残りの主張も同様の方法で証明できる．詳細は読者に委ねる (問題 13.4)．■

本章の最後に正定値，非負定値の概念を解説しておく．

**定義 13.12**　$A = (a_{ij}) \in \boldsymbol{K}^{N \times N}$ をエルミート行列とする．エルミート形式 $Q_A$ が**正定値**であるとは，任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{x} \neq \mathbf{0}$ に対して,

$$Q_A(\mathbf{x}) > 0$$

をみたすことである．また，任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{x} \neq \mathbf{0}$ に対して,

$$Q_A(\mathbf{x}) \geq 0$$

をみたすものを**半正定値**あるいは**非負定値**であるという．またエルミート行列 $A$ が正定値であるとは，$Q_A$ が正定値のことであり，$A$ が非負定値とは $Q_A$ が非負定値のことである．

**例 13.7**　$I_N$ は正定値である．また

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

は半正定値である．

**例 13.8**　$B \in \boldsymbol{K}^{N\times M}$ とする．このとき次のことが成り立つ．

(1)　$BB^* \in \boldsymbol{K}^{N\times N}$, $B^*B \in \boldsymbol{K}^{M\times M}$ はともにエルミート行列であり，非負定値である．

(2)　$BB^*$ が正定値であるための必要十分条件は $\mathrm{rank}(B) = N$ となることである．

(3)　$B^*B$ が正定値であるための必要十分条件は $\mathrm{rank}\,(B) = M$ となることである．

**証明.** $(BB^*)^* = BB^*$ よりエルミート行列である．任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x}^*BB^*\mathbf{x} = (B^*\mathbf{x})^*\,(B^*\mathbf{x}) = \|B^*\mathbf{x}\|^2 \geq 0$$

である．ゆえに非負定値である．後半の主張を証明する．$\mathrm{rank}\,(B) = r$ とすると $\mathrm{rank}\,(B^*) = r$ であり

$$\dim \mathcal{N}\,(B^*) = N - \dim \mathcal{R}\,(B^*) = N - r$$

が成り立つ．もし $r < N$ ならば，$\dim \mathcal{N}\,(B^*) \geq 1$ より $\mathbf{x} \neq \mathbf{0}$ かつ $\mathbf{x}^*BB^*\mathbf{x} = \|B^*\mathbf{x}\|^2 = 0$ なるものが存在するから，$B$ は正定値ではない．もし $r = N$ ならば $\dim \mathcal{N}\,(B^*) = 0$ であるから，任意の $\mathbf{x} \neq \mathbf{0}$ に対して $\|B^*\mathbf{x}\| > 0$ となる．

$A = B^* \in \boldsymbol{K}^{M\times N}$ とおく．上記の議論を $A$ に適用することにより，$AA^*$ $(= B^*B$ が非負定値であることがわかる．また $AA^*$ $(= B^*B)$ が正定値であるための必要十分条件は $\mathrm{rank}\,(A) = M$ となることがわかるが，

$$\mathrm{rank}\,(A) = \mathrm{rank}\,(B^*) = \mathrm{rank}\,(B)$$

である．■

正定値行列のもつ性質として次のようなことがあげられる．

**命題 13.28**　(1)　正定値行列は正則である．

(2)　非負定値行列の対角成分は非負，また正定値行列の対角成分は正である．

**証明.** (1)　$A \in \boldsymbol{K}^{N\times N}$ を正定値行列とする．$\mathbf{x} \in \mathcal{N}(A)$ とする．このとき，$\mathbf{x}^*A\mathbf{x} = 0$ であるから，正定値の定義から $\mathbf{x} = \mathbf{0}$ でなければならない．したがって $A$ は正則である．

(2)　$A = (a_{jk})$ とすると，$a_{kk} = \mathbf{e}_k^{(N)*}A\mathbf{e}_k^{(N)}$ である．■

正定値，半正定値であることは $A$ の固有値によって次のように特徴付けられる．

**定理 13.29**　$A \in \boldsymbol{K}^{N\times N}$ をエルミート行列とする．$Q_A$ が正定値であるための必要十分条件は，$A$ のすべての固有値が正となることである．また $Q_A$ が非負定値であるための必要十分条件は，$A$ のすべての固有値が非負になることである．

**証明.** $A$ はあるユニタリー行列 $U$ により対角化できる．任意の $\mathbf{y} = (y_j) \in \boldsymbol{K}^N$ に対して，$\mathbf{x} = U\mathbf{y}$ とし，$\lambda_1, \cdots, \lambda_N$ を $A$ の固有値とする．このとき

$$\begin{aligned} Q_A(\mathbf{x}) &= \mathbf{x}^* A\mathbf{x} = (U\mathbf{y})^* AU\mathbf{y} = \mathbf{y}^* \operatorname{diag}_{N,N}(\lambda_1, \cdots, \lambda_N)\,\mathbf{y} \\ &= \lambda_1 |y_1|^2 + \cdots + \lambda_N |y_N|^2 \end{aligned}$$

と表される．$\mathbf{x} \neq \mathbf{0}$ ならば $\mathbf{y} \neq \mathbf{0}$ であることに注意すれば，容易に定理が証明される．■

この定理から容易に次のことが示せる．

**系 13.30**　$A \in \boldsymbol{K}^{N\times N}$ が非負定値であるとする．このとき $\det A \geq 0$ である．また $A$ が正定値ならば $\det A > 0$ である．

**証明.** $\lambda_1, \cdots, \lambda_N$ を $A$ の固有値とする．あるユニタリー行列により，$A = U\operatorname{diag}(\lambda_1, \cdots, \lambda_N)\, U^{-1}$ と表せているので，

$$\det A = \lambda_1 \cdots \lambda_N$$

であるから，定理 13.29 より定理が示される．■

**問題 13.4**　定理 13.27 のうち，

$$\lambda_j = \max_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B\cap V} Q_A(\mathbf{x})$$

を証明せよ．

**問題 13.5**　$A \in \boldsymbol{K}^{N\times N}$ が正定値であるための必要十分条件は，ある $c > 0$ が存在し，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x}^* A\mathbf{x} \geq c\,\|\mathbf{x}\|^2$$

が成り立つことである．(ヒント：必要なら，有界閉集合上の連続関数は最小値をもつという事実を用いよ．)

**問題 13.6**　定理 13.26 において $A$ が非負定値かつ $\lambda_N > 0$ とする．$\mathbf{x} \in B$ が

$Q_A(\mathbf{x}) = \lambda_N$ をみたせば $\mathbf{x}$ は $\lambda_N$ の固有ベクトルであることを示せ.

## 13.5 幾何的な問題と主成分分析への応用

いま $\boldsymbol{R}^2$ 上に $N$ 個の点

$$\mathbf{x}_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}, \quad \cdots, \quad \mathbf{x}_N = \begin{pmatrix} x_{N1} \\ x_{N2} \end{pmatrix}$$

が与えられているとする．$\mathbf{w} \in \boldsymbol{R}^2$, $\|\mathbf{w}\| = 1$ とし，$\mathbf{a} \in \boldsymbol{R}^2$, $\langle \mathbf{w}, \mathbf{a} \rangle = 0$ とする．

$$l : \mathbf{x} = \mathbf{a} + t\mathbf{w}$$

なる直線を考える (図 13.2 左図)．点 $\mathbf{x}_j$ と直線 $l$ の点との最短距離を $\delta_j$ により表す．すなわち

$$\delta_j = \min \{ \|\mathbf{x}_j - \mathbf{x}\| : \mathbf{x} \text{ は } l \text{ 上の点} \}$$

である．このとき

$$\delta = \sqrt{\sum_{j=1}^{N} \delta_j^2} \tag{13.16}$$

とすると $\delta$ は直線のとり方によりその値が変化する (図 13.2 参照).

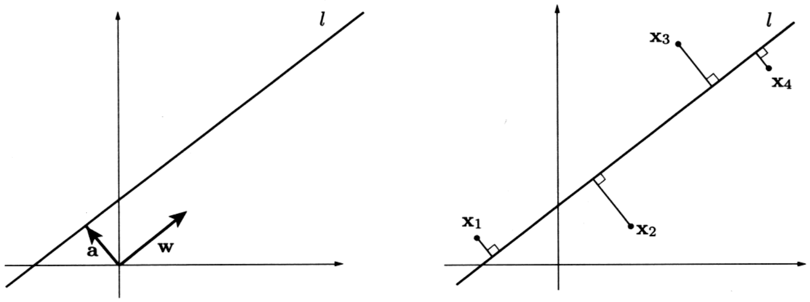



図 13.2

本節では次のような問題を考察する．

| **問題 P**　$\delta$ が最も小さくなるような直線 $l$ を求めよ． |
|---|

まずいくつかの準備的な考察からはじめる．

**補題 13.31**　点 $\mathbf{x} \in \boldsymbol{R}^2$ から直線 $l$ に下ろした垂線と直線 $l$ との交点を表すベクトル $\mathbf{y}$ は

$$\mathbf{y} = \langle \mathbf{x}, \mathbf{w} \rangle \, \mathbf{w} + \mathbf{a}$$

である．

**証明．** $\mathbf{x}' = \mathbf{x} - \mathbf{a}$ とし，$l' : \mathbf{x} = t\mathbf{w}$ とする．$\mathbf{x}'$ から $l'$ に下ろした垂線と $l'$ との交点 $\mathbf{y}'$ は $\mathbf{x}'$ の $l'$ への直交射影になっている．それは

$$\mathbf{y}' = \langle \mathbf{x}', \mathbf{w} \rangle \, \mathbf{w}$$

で与えられる．したがって，

$$\begin{aligned}\mathbf{y} &= \mathbf{y}' + \mathbf{a} = \langle \mathbf{x}', \mathbf{w} \rangle \, \mathbf{w} + \mathbf{a} = \langle \mathbf{x} - \mathbf{a}, \mathbf{w} \rangle \, \mathbf{w} + \mathbf{a} \\ &= \langle \mathbf{x}, \mathbf{w} \rangle \, \mathbf{w} + \mathbf{a}.\end{aligned}$$

■

このことを用いて次のことが証明される．

**補題 13.32**

$$\overline{x}_k = \frac{1}{N} \sum_{j=1}^{N} x_{jk}, \quad k = 1, 2$$

とし，

$$\overline{\mathbf{x}} = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} = \frac{1}{N} \left( \mathbf{x}_1, \cdots, \mathbf{x}_N \right) \mathbf{1}_{N \times 1}$$

とおく（$\mathbf{1}_{N \times 1}$ は (2.10) 参照）．$\delta$ が最も小さくなるような直線は必ず点 $\overline{\mathbf{x}}$ を通る．

**証明．** $\mathbf{y}_j = \langle \mathbf{x}_j, \mathbf{w} \rangle \, \mathbf{w} + \mathbf{a}$ とすると，$\delta_j = \| \mathbf{x}_j - \mathbf{y}_j \|$ であるから，

$$\begin{aligned}\delta^2 &= \sum_{j=1}^{N} \| \mathbf{x}_j - \mathbf{y}_j \|^2 = \sum_{j=1}^{N} \| \mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \, \mathbf{w} - \mathbf{a} \|^2 \\ &= \sum_{j=1}^{N} \| \mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \, \mathbf{w} - \overline{\mathbf{x}} - (\mathbf{a} - \overline{\mathbf{x}}) \|^2 \\ &= \sum_{j=1}^{N} \| \mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \, \mathbf{w} - \overline{\mathbf{x}} \|^2 \\ &\quad - 2 \sum_{j=1}^{N} \langle \mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \, \mathbf{w} - \overline{\mathbf{x}}, \mathbf{a} - \overline{\mathbf{x}} \rangle + N \| \mathbf{a} - \overline{\mathbf{x}} \|^2\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{N} \|\mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \mathbf{w} - \overline{\mathbf{x}}\|^2 \\
&\quad - 2 \langle N\overline{\mathbf{x}} - \langle N\overline{\mathbf{x}}, \mathbf{w} \rangle \mathbf{w} - N\overline{\mathbf{x}}, \mathbf{a} - \overline{\mathbf{x}} \rangle + N \|\mathbf{a} - \overline{\mathbf{x}}\|^2 \\
&= \sum_{j=1}^{N} \|\mathbf{x}_j - \langle \mathbf{x}_j, \mathbf{w} \rangle \mathbf{w} - \overline{\mathbf{x}}\|^2 - 2N \langle \overline{\mathbf{x}}, \mathbf{w} \rangle^2 + N \|\mathbf{a} - \overline{\mathbf{x}}\|^2 .
\end{aligned}
$$

ここで最初の 2 項は $\mathbf{a}$ に関係ないから，$\mathbf{a} = \overline{\mathbf{x}}$ の場合に $\delta$ が最小になる． ■

$\overline{\mathbf{x}}$ を $\mathbf{x}_1, \cdots, \mathbf{x}_N$ の平均ベクトルという．

さて，

$$
\widetilde{X} = \begin{pmatrix} x_{11} - \overline{x}_1 & x_{12} - \overline{x}_2 \\ \vdots & \vdots \\ x_{N1} - \overline{x}_1 & x_{N2} - \overline{x}_2 \end{pmatrix} = \begin{pmatrix} \widetilde{x}_{11} & \widetilde{x}_{12} \\ \vdots & \vdots \\ \widetilde{x}_{N1} & \widetilde{x}_{N2} \end{pmatrix}
$$

とし，

$$
S_X = \frac{1}{N} \widetilde{X}^T \widetilde{X}
$$

とおく．このとき，次のことが成り立つ．

**定理 13.33** $S_X \neq O$, $\mathbf{w} \in \boldsymbol{R}^2$, $\|\mathbf{w}\| = 1$ とする．このとき，次の (1), (2) は同値である．

(1) 方程式

$$
\overline{\mathbf{x}} + t\mathbf{w}
$$

で表される直線は，$\delta$ が最も小さくなるような直線である．

(2) $\mathbf{w}$ は $S_X$ の最大の固有値に属する長さ 1 の固有ベクトルである．

この定理は次の補題から証明される．

**補題 13.34** $B = \{\mathbf{x} \in \boldsymbol{R}^2 : \|\mathbf{x}\| = 1\}$ とおく．このとき

$$
\min \delta^2 = N \left\{ \mathrm{tr}(S_X) - \max_{\mathbf{w} \in B} Q_{S_X}(\mathbf{w}) \right\}
$$

である．

**証明.** 直線 $l : \mathbf{x} = \overline{\mathbf{x}} + t\mathbf{w}$ を $-\overline{\mathbf{x}}$ 平行移動した直線は $l' : \mathbf{x} = t\mathbf{w}$ である．

$$
\widetilde{\mathbf{x}}_j = \mathbf{x}_j - \overline{\mathbf{x}} = \begin{pmatrix} \widetilde{x}_{j1} \\ \widetilde{x}_{j2} \end{pmatrix}
$$

とおく．$\widetilde{\mathbf{x}}_j$ から $l'$ に下ろした垂線と $l'$ の交点は，$\widetilde{\mathbf{x}_j}$ の $\mathrm{span}\,\{\mathbf{w}\}$ への直交射影であるから

$$\delta^2 = \sum_{j=1}^{N} \|\widetilde{\mathbf{x}}_j - \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}\|^2$$

である．ここで，$\widetilde{\mathbf{x}}_j - \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}$ と $\langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}$ が直交していることに注意すれば，

$$\begin{aligned}\|\widetilde{\mathbf{x}}_j\|^2 &= \|\widetilde{\mathbf{x}}_j - \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w} + \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}\|^2 \\ &= \|\widetilde{\mathbf{x}}_j - \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}\|^2 + \|\langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}\|^2 \\ &= \|\widetilde{\mathbf{x}}_j - \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle \mathbf{w}\|^2 + \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle^2\end{aligned}$$

が成り立っていることがわかる．したがって

$$\delta^2 = \sum_{j=1}^{N} \|\widetilde{\mathbf{x}}_j\|^2 - \sum_{j=1}^{N} \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle^2 . \tag{13.17}$$

さて，

$$\begin{aligned}Q_{S_X}(\mathbf{w}) &= \mathbf{w}^T S_X \mathbf{w} \\ &= \frac{1}{N}\begin{pmatrix} w_1 & w_2 \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{N} \widetilde{x}_{j1}\widetilde{x}_{j1} & \sum_{j=1}^{N} \widetilde{x}_{j1}\widetilde{x}_{j2} \\ \sum_{j=1}^{N} \widetilde{x}_{j2}\widetilde{x}_{j1} & \sum_{j=1}^{N} \widetilde{x}_{j2}\widetilde{x}_{j2} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ &= \frac{1}{N}\sum_{j=1}^{N} (\widetilde{x}_{j1}w_1\widetilde{x}_{j1}w_1 + 2\widetilde{x}_{j1}w_1\widetilde{x}_{j2}w_2 + \widetilde{x}_{j2}w_2\widetilde{x}_{j2}w_2) \\ &= \frac{1}{N}\sum_{j=1}^{N} (\widetilde{x}_{j1}w_1 + \widetilde{x}_{j2}w_2)^2 = \frac{1}{N}\sum_{j=1}^{N} \langle \widetilde{\mathbf{x}}_j, \mathbf{w}\rangle^2\end{aligned}$$

であり，

$$\mathrm{tr}(S_X) = \frac{1}{N}\sum_{j=1}^{N} \left((\widetilde{x}_{j1})^2 + (\widetilde{x}_{j2})^2\right) = \frac{1}{N}\sum_{j=1}^{N} \|\widetilde{\mathbf{x}}_j\|^2$$

より，

$$\delta^2 = N\left\{\mathrm{tr}(S_X) - Q_{S_X}(\mathbf{w})\right\}$$

である．これより補題が示される．■

**定理 13.33 の証明．** 定理 13.26 より，$\max_{\mathbf{w}\in B} Q_{S_X}(\mathbf{w})$ は $S_X$ の最大固有値であり，それを与える $\mathbf{w}$ は最大固有値に属する長さ 1 の固有ベクトルである (問題 13.6

参照)．よって定理が証明された． ■

**多変量解析における解釈**

問題 P と定理 13.33 は多変量解析に応用がある．多変量解析との関連を見ておこう．なお，多変量解析については後述の第 17 章も参照してほしい．多変量解析では

$$X = \begin{pmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{pmatrix}$$

とおき，これを (2 変量の) **データ行列**という．たとえば，あるクラスで行った数学と英語の試験の点数が

| 学生番号 | 数学 | 英語 |
|---|---|---|
| 1 | 78 | 82 |
| 2 | 55 | 30 |
| ⋮ | ⋮ | ⋮ |
| 60 | 68 | 75 |

(13.18)

であるとする．これを $N = 60$ とし

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{pmatrix} = \begin{pmatrix} 78 & 82 \\ 55 & 30 \\ \vdots & \vdots \\ 68 & 75 \end{pmatrix}$$

としたものがデータ行列である．

$$\overline{\mathbf{x}} = \left(\frac{1}{N}\mathbf{1}_{1\times N}X\right)^T = \begin{pmatrix} N^{-1}\displaystyle\sum_{j=1}^{N} x_{j1} \\ N^{-1}\displaystyle\sum_{j=1}^{N} x_{j2} \end{pmatrix}$$

とし，これをデータ行列 $X$ に対する**平均ベクトル**という．また

$$\widetilde{X} = (\widetilde{x}_{ij}) = X - \mathbf{1}_{N\times 1}\overline{\mathbf{x}}^T$$

は**平均偏差行列**，$S_X = N^{-1}\widetilde{X}^T\widetilde{X}$ は**分散共分散行列**と呼ばれている．

$$S_X = \frac{1}{N}\begin{pmatrix} x_{11}-\overline{x}_1 & \cdots & x_{N1}-\overline{x}_1 \\ x_{12}-\overline{x}_2 & \cdots & x_{N2}-\overline{x}_2 \end{pmatrix}\begin{pmatrix} x_{11}-\overline{x}_1 & x_{12}-\overline{x}_2 \\ \vdots & \vdots \\ x_{N1}-\overline{x}_1 & x_{N2}-\overline{x}_2 \end{pmatrix}$$

$$= \frac{1}{N}\begin{pmatrix} \sum_{j=1}^{N}(x_{j1}-\overline{x}_1)(x_{j1}-\overline{x}_1) & \sum_{j=1}^{N}(x_{j1}-\overline{x}_1)(x_{j2}-\overline{x}_2) \\ \sum_{j=1}^{N}(x_{j2}-\overline{x}_2)(x_{j1}-\overline{x}_1) & \sum_{j=1}^{N}(x_{j2}-\overline{x}_2)(x_{j2}-\overline{x}_2) \end{pmatrix}$$

となっている．

さて，たとえば (13.18) の試験結果から，序列をつけるなど評価をする必要があるとしよう．そのため数学と英語の得点の平均偏差に何らかの重みをつけて，

$$f_n = w_1\widetilde{x}_{n1} + w_2\widetilde{x}_{n2}, \quad n = 1, \cdots, N$$

を考え，そして点数 $f_n$ によって評価することを考えてみよう．ただし重みは $\mathbf{w}^T = (w_1, w_2)$, $\|\mathbf{w}\| = 1$ となるようにとる．さて，重みをつけて個々人の得点の傾向の差異を分析するためには，なるべく分散

$$\sum_{n=1}^{N}\left(f_n - \overline{f}\right)^2$$

(ただし $\overline{f} = \frac{1}{N}\sum_{n=1}^{N} f_n$) が大きい方がよい．そこで分散 $\sum_{n=1}^{N}\left(f_n - \overline{f}\right)^2$ が最大になるような重み $\mathbf{w}$ を見つけ，それに対して $f_1, \cdots, f_N$ を計算する．この重みをつけて得られた得点を第 1 主成分得点という．個人の得点の傾向をこのようにして分析する方法を主成分分析という．じつはこの分散を最大にする重み $\mathbf{w}$ をもとめることと，問題 P を解くことが同値になっているのである．このことを証明しよう．いま

$$\overline{f} = \frac{1}{N}\sum_{n=1}^{N} f_n = \frac{1}{N}\sum_{n=1}^{N}(w_1\widetilde{x}_{n1} + w_2\widetilde{x}_{n2})$$

$$= \frac{1}{N}\left(w_1\sum_{n=1}^{N}\widetilde{x}_{n1} + w_2\sum_{n=1}^{N}\widetilde{x}_{n2}\right) = \frac{1}{N}(w_1 0 + w_2 0) = 0$$

であるから

$$\sum_{n=1}^{N}\left(f_n - \overline{f}\right)^2 = \sum_{n=1}^{N} {f_n}^2 = \sum_{n=1}^{N}(w_1\widetilde{x}_{n1} + w_2\widetilde{x}_{n2})^2$$

$$
\begin{aligned}
&= \sum_{n=1}^{N} \left( w_1^2 \widetilde{x}_{n1}^2 + 2 w_1 \widetilde{x}_{n1} w_2 \widetilde{x}_{n2} + w_2^2 \widetilde{x}_{n2}^2 \right) \\
&= \begin{pmatrix} w_1 & w_2 \end{pmatrix} \begin{pmatrix} \displaystyle\sum_{n=1}^{N} \widetilde{x}_{n1}^2 & \displaystyle\sum_{n=1}^{N} \widetilde{x}_{n1}\widetilde{x}_{n2} \\ \displaystyle\sum_{n=1}^{N} \widetilde{x}_{n1}\widetilde{x}_{n2} & \displaystyle\sum_{n=1}^{N} \widetilde{x}_{n2}^2 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\
&= N Q_{S_X}(\mathbf{w})
\end{aligned}
$$

が得られる．したがって分散を最大にする長さ 1 の重み $\mathbf{w}$ は，$Q_{S_X}(\mathbf{w})$ を最大にする長さ 1 のベクトルである．それは $S_X$ の最大固有値に属する長さ 1 のベクトルであり，また補題 13.34 より，$\delta$ が最も小さくなるような直線の方程式を与える $\mathbf{w}$ に他ならない．

このように分散 $\sum_{n=1}^{N} \left( f_n - \overline{f} \right)^2$ が最大になることは問題 P と同値になっている．

ところで問題 P そのものの多変量解析における解釈はどのようなものであろうか？　このことを説明するために，主成分分析におけるデータのもつ情報量の定義をしておこう．データ行列 $X = (x_{kp}) \in \boldsymbol{R}^{N \times 2}$ に対して $X(k, :)$ を個体 $k$ のデータといい，

$$
(x_{k1} - \overline{x}_1)^2 + (x_{k2} - \overline{x}_2)^2
$$

を個体 $k$ のもつ情報量という．また

$$
\sum_{k=1}^{N} \left\{ (x_{k1} - \overline{x}_1)^2 + (x_{k2} - \overline{x}_2)^2 \right\}
$$

をデータ全体の情報量という．(13.17) より

$$
\sum_{k=1}^{N} \left\{ (x_{k1} - \overline{x}_1)^2 + (x_{k2} - \overline{x}_2)^2 \right\} = \sum_{k=1}^{N} f_k^2 + \delta^2
$$

となっている．ここで $f_k^2 = \left( f_k - \overline{f} \right)^2$ であるから，$f_k^2$ は個体 $k$ の第 1 主成分得点に関する情報量とみなせる．したがって，$f_k^2$ は第 1 主成分得点が吸収できる情報量を表し，$\delta^2$ は吸収できない情報量を表す．問題 P を求めるということは，できるだけ吸収できない情報量の総和を小さく抑えるような直線を求めることに相当する．

主成分分析の例とより詳細な内容については第 17.3 節で述べる．

# 第14章　行列の分解定理

この章では，線形代数の応用においてよく使われるいくつかの分解定理を述べる．

**ポイント 14**　本章で述べた行列の分解は，数値計算などで使われる．

## 14.1　*LU* 分解

$LU$ 分解は連立 1 次方程式の解を効率よく求めるために使われる分解で，次のように定義される．

**定義 14.1**　$(N, N)$ 行列 $A$ がある上三角行列 $U \in \boldsymbol{K}^{N\times N}$ とある単位下三角行列 $L \in \boldsymbol{K}^{N\times N}$ とによって

$$A = LU \tag{14.1}$$

と分解できるとき，$A$ は $LU$ **分解可能**であるといい，(14.1) を $A$ の $LU$ **分解**という．

もし $A$ が $LU$ 分解可能で，さらに $L, U$ の対角成分が 0 でない場合，方程式

$$A\mathbf{x} = \mathbf{b}$$

は次のように前進代入法と後退代入法のみを使って容易に解くことができる．まず方程式

$$L\mathbf{x}' = \mathbf{b}$$

の解を前進代入法で計算する．次に方程式

$$U\mathbf{x} = \mathbf{x}'$$

の解を後退代入法で計算する．これにより，

$$A\mathbf{x} = L(U\mathbf{x}) = L\mathbf{x}' = \mathbf{b}$$

が導かれる．したがって，$\mathbf{x}$ は $A\mathbf{x} = \mathbf{b}$ の解になっている．

このように $LU$ 分解しておけば，与えられたデータ $\mathbf{b}$ がいろいろ変わっても，データごとにガウスの消去法を用いて計算する必要がなくなる．

また行列 $A$ の行列式も，$A$ が $LU$ 分解されていれば，$U$ の対角成分 $u_{jj}$, $1 \leq j \leq N$ により

$$\det A = u_{11}u_{22}\cdots u_{NN}$$

と求めることも可能である (例題 5.2 参照)．

以下では具体的な $LU$ 分解の構成方法を述べる．この分解はすでに学んだガウスの消去法を用いて求めることができる．ガウスの消去法の方法を復習しながらこのことを示す．$A = (a_{ij}) \in \boldsymbol{K}^{N\times N}$ を正則な行列とする．

まず $A$ の第 1 列の成分がすべて 0 であるならば $A$ は正則でないから，$A$ の第 1 列の成分のいずれかは 0 ではない．次の二つの場合が考えられる．

場合 1-1 : $a_{11} \neq 0$.
場合 1-2 : $a_{11} = 0$.

もし場合 1-2 が起こった場合は，$a_{k1} \neq 0$ であるような $k > 1$ の一つを選び第 1 行と第 $k$ 行と入れ替える．たとえば $|a_{k1}| = \max_{2\leq m\leq N} |a_{m1}|$ をみたすような $k$ をとる．この入れ替えの操作は基本行列 $F_{1k}$ を左からかけることにより実行される．

場合 1-1 の場合は $P_1 = I_N$ とし，そして場合 1-2 の場合は $P_1 = F_{1k}$ とおく．いずれの場合にも $P_1A = \left(a_{ij}^{(1)}\right)$ とおく．このとき $a_{11}^{(1)} \neq 0$ である．そこで

$$E_{N1}\left(-\frac{a_{N1}^{(1)}}{a_{11}^{(1)}}\right)\cdots E_{31}\left(-\frac{a_{31}^{(1)}}{a_{11}^{(1)}}\right)E_{21}\left(-\frac{a_{21}^{(1)}}{a_{11}^{(1)}}\right)P_1A$$

と行に関する基本変形をする．

$$M_1 = E_{N1}\left(-\frac{a_{N1}^{(1)}}{a_{11}^{(1)}}\right)\cdots E_{31}\left(-\frac{a_{31}^{(1)}}{a_{11}^{(1)}}\right)E_{21}\left(-\frac{a_{21}^{(1)}}{a_{11}^{(1)}}\right)$$

とおくと，簡単な計算で

$$M_1 = \begin{pmatrix} 1 & & & & \\ -\dfrac{a_{21}^{(1)}}{a_{11}^{(1)}} & \ddots & & O & \\ -\dfrac{a_{31}^{(1)}}{a_{11}^{(1)}} & 0 & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ -\dfrac{a_{N1}^{(1)}}{a_{11}^{(1)}} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

となっていることがわかる．$A^{(1)\prime} = M_1 P_1 A$ とおく．これを計算した結果は

$$A^{(1)\prime} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(1)\prime} & \cdots & a_{2N}^{(1)\prime} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{N2}^{(1)\prime} & \cdots & a_{NN}^{(1)\prime} \end{pmatrix}$$

の形の行列になる．

次に $A^{(1)\prime}$ の第 2 列をみる．もし $a_{j2}^{(1)\prime} = 0, j \geq 2$ であれば，$\operatorname{rank}(A) = \operatorname{rank}\left(A^{(1)\prime}\right) \leq N - 1$ となって $A$ が正則であることに反するので，いずれかの $a_{j2}^{(1)\prime}, j \geq 2$ は 0 ではない．次の二つの場合が考えられる．

場合 2-1 : $a_{22}^{(1)\prime} \neq 0.$
場合 2-2 : $a_{22}^{(1)\prime} = 0.$

もし場合 2-2 が起こった場合は，$a_{k2}^{(1)\prime} \neq 0$ であるような $k > 2$ の一つを選び第 2 行と第 $k$ 行とを入れ替える．入れ替える操作は基本行列 $F_{2k}$ を左からかけることにより実行される．場合 2-1 の場合は $P_2 = I_N$，そして場合 2-2 の場合は $P_2 = F_{2k}$ とおき，どちらの場合にも $P_2 A^{(1)\prime} = \left(a_{ij}^{(2)}\right)$ とおく．このとき $a_{22}^{(2)} \neq 0$ である．そこで

$$E_{N2}\left(-\frac{a_{N2}^{(2)}}{a_{22}^{(2)}}\right) \cdots E_{32}\left(-\frac{a_{32}^{(2)}}{a_{22}^{(2)}}\right) P_2 A^{(1)\prime}$$

を計算する．

$$M_2 = E_{N2}\left(-\frac{a_{N2}^{(2)}}{a_{22}^{(2)}}\right) \cdots E_{32}\left(-\frac{a_{32}^{(2)}}{a_{22}^{(2)}}\right)$$

とおくと

$$M_2 = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & O & \\ & -\dfrac{a_{32}^{(2)}}{a_{22}^{(2)}} & 1 & & & \\ \vdots & -\dfrac{a_{42}^{(2)}}{a_{22}^{(2)}} & 0 & \ddots & & \\ & \vdots & \vdots & \ddots & 1 & \\ 0 & -\dfrac{a_{N2}^{(2)}}{a_{22}^{(2)}} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

である．

$$A^{(2)\prime} = M_2 P_2 A^{(1)\prime} = M_2 P_2 M_1 P_1 A$$

とする．計算により

$$A^{(2)\prime} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2N}^{(2)} \\ 0 & 0 & a_{33}^{(2)\prime} & \cdots & a_{3N}^{(2)\prime} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{N3}^{(2)\prime} & \cdots & a_{NN}^{(2)\prime} \end{pmatrix}$$

が示せる．以下この操作を続けていくと

$$M_1,\, M_2,\, \cdots,\, M_{N-1},\, P_1,\, P_2,\, \cdots,\, P_{N-1} \tag{14.2}$$

および

$$A^{(N-1)\prime} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & a_{1N}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & \cdots & a_{2N}^{(2)} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & a_{N-1,N-1}^{(N-1)} & a_{N-1,N}^{(N-1)} \\ 0 & 0 & \cdots & 0 & a_{NN}^{(N)\prime} \end{pmatrix}$$

が得られる．そこで $U = A^{(N-1)\prime}$ とおく．

$$U = M_{N-1} P_{N-1} \cdots M_1 P_1 A$$

が成り立っている．

さて，いま条件

$$P_1 = \cdots = P_{N-1} = I_N \tag{14.3}$$

が成り立っているとしよう．このとき

$$U = M_{N-1} \cdots M_1 A$$

であるが，$E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda)$ であることに注意すると，

$$M_j^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & \dfrac{a_{j+1,j}^{(j)}}{a_{j,j}^{(j)}} & 1 & & & \\ O & & \dfrac{a_{j+2,j}^{(j)}}{a_{j,j}^{(j)}} & 0 & \ddots & & \\ & & \vdots & \vdots & \ddots & \ddots & \\ & & \dfrac{a_{N,j}^{(j)}}{a_{j,j}^{(j)}} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

となることがわかる．$L = M_1^{-1} \cdots M_{N-1}^{-1}$ とすると

$$L = \begin{pmatrix} 1 & & & & \\ \dfrac{a_{21}^{(1)}}{a_{11}^{(1)}} & 1 & & O & \\ \dfrac{a_{31}^{(1)}}{a_{11}^{(1)}} & \dfrac{a_{32}^{(2)}}{a_{22}^{(2)}} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \dfrac{a_{N1}^{(1)}}{a_{11}^{(1)}} & \dfrac{a_{N2}^{(2)}}{a_{22}^{(2)}} & \cdots & \dfrac{a_{N,N-1}^{(N-1)}}{a_{N-1,N-1}^{(N-1)}} & 1 \end{pmatrix}$$

であり，

$$A = LU$$

が成り立っている．すなわち (14.3) が成り立っていれば，$A$ は $LU$ 分解できることがわかる．

さらに $LU$ 分解できるための条件について議論を進める．その前に，単位下三角行列と正則な上三角行列による $LU$ 分解の一意性を示しておく．

**定理 14.1** $A \in \boldsymbol{K}^{N\times N}$ が単位下三角行列 $L_1, L_2$ および正則な上三角行列 $U_1, U_2$ により

$$A = L_1U_1 = L_2U_2$$

と分解できるならば，$L_1 = L_2$, $U_1 = U_2$ である．

**証明.** $L_1U_1 = L_2U_2$ より

$$L_2^{-1}L_1 = U_2U_1^{-1}$$

である．問題 2.3 (2) と定理 8.4 より $L_2^{-1}L_1$ は単位下三角行列，$U_2^{-1}U_1$ は正則な上三角行列であることがわかるから，$L_2^{-1}L_1$ は単位下三角行列であり，かつ上三角行列である．ゆえに $L_2^{-1}L_1 = I_N$ でなければならない．それゆえ $L_1 = L_2$, $U_2 = U_1$ が得られる． ■

条件 (14.3) があれば $LU$ 分解できることを示したが，さらにこれを用いて $LU$ 分解できるための一つの必要十分条件を求めることができる．

**定理 14.2** $A \in \boldsymbol{K}^{N\times N}$ が単位下三角行列 $L$ と正則な上三角行列 $U$ により $LU$ 分解可能であるための必要十分条件は，

$$\det A\,(1:k, 1:k) \neq 0, \quad k = 1, \cdots, N \tag{14.4}$$

が成り立つことである．

さらにこのとき，$U = (u_{ij})$ とおくと，

$$\begin{aligned} u_{11} &= \det A\,(1:1, 1:1) \\ u_{kk} &= \frac{\det A\,(1:k, 1:k)}{\det A\,(1:k-1, 1:k-1)}, \quad k \geq 2 \end{aligned} \tag{14.5}$$

が成り立っている．

**証明.** まず十分性を示す．$A = (a_{mn})$ とする．条件より $a_{11} \neq 0$ である．したがって $P_1 = I_N$ とし, $M_1$ が定義できる．帰納法で示す．もし $P_1 = \cdots = P_{k-1} = I_N$ かつ $M_{k-1}$ まで定義できたとすると，このとき，

$$A = M_1^{-1} \cdots M_{k-1}^{-1} \begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1,k}^{(1)} & \cdots & a_{1,N}^{(1)} \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & a_{k,k}^{(k)\prime} & \cdots & a_{k,N}^{(k)\prime} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{N,k}^{(k)\prime} & \cdots & a_{N,N}^{(k)\prime} \end{pmatrix}$$

である．右辺を次のようにブロック行列に区分けする．ただし次の式において，$L_{kk}$ は単位下三角行列，$U_{kk}$ は上三角行列である．

$$\begin{aligned} A &= \begin{pmatrix} L_{kk} & O_{k,N-k} \\ K_{N-k,k} & I_{N-k,N-k} \end{pmatrix} \begin{pmatrix} U_{kk} & B_{k,N-k} \\ C_{N-k,k} & D_{N-k,N-k} \end{pmatrix} \\ &= \begin{pmatrix} L_{kk}U_{kk} & * \\ * & * \end{pmatrix} \end{aligned}$$

(ここで $*$ は適切な行列を表している)．このとき $L_{kk}$ の対角成分は 1 であり，$\det U_{kk} = a_{11}^{(1)} \cdots a_{k-1,k-1}^{(k-1)} a_{kk}^{(k)\prime}$ であるから，

$$0 \neq \det A\,(1:k, 1:k) = a_{11}^{(1)} \cdots a_{k-1,\,k-1}^{(k-1)} a_{kk}^{(k)\prime}$$

となっている．ゆえに $a_{kk}^{(k)\prime} \neq 0$ であり，$P_k = I_N$ ととることができ，$M_k$ が定義できる．結局，帰納法により $LU$ 分解可能であることが証明された．以上の議論から $U$ の正則性は明らかである．

次に必要性を示す．いま $A = LU$ とする．$U = (u_{ij})$ とおく．$U$ は正則であるから

$$u_{11} \cdots u_{NN} = \det U \neq 0$$

である．$L_k = L\,(1:k, 1:k)$, $U_k = U\,(1:k, 1:k)$ とおき，$L, U$ をブロック行列で表すと

$$A = \begin{pmatrix} L_k & O \\ B & R \end{pmatrix} \begin{pmatrix} U_k & C \\ O & Q \end{pmatrix} = \begin{pmatrix} L_kU_k & * \\ * & * \end{pmatrix}$$

となっている．ただし，ここで $*$ は適切なブロック行列を表す．したがって，定理 14.3 より

$$\det A\,(1:k, 1:k) = \det L_k \det U_k = u_{11} \cdots u_{kk} \neq 0$$

が成り立つ．このことから後半の主張も得られる．■

条件 (14.4) をみたす例として次のものがある．

**定理 14.3** $A \in \boldsymbol{K}^{N\times N}$ を正定値行列とする．このとき，任意の $k=1,\cdots,N$ に対して

$$\det A\,(1:k,1:k) > 0$$

が成り立っている．

**証明．** $X = \left(\mathbf{e}_1^{(N)},\cdots,\mathbf{e}_k^{(N)}\right) \in \boldsymbol{K}^{N\times k}$ とする．このとき

$$A(1:k,1:k) = X^* A X$$

である．$\boldsymbol{K}^k \ni \mathbf{z} \neq \mathbf{0}$ とすると $X\mathbf{z} = \begin{pmatrix} \mathbf{z} \\ \mathbf{0}_{N-k} \end{pmatrix} \neq \mathbf{0}_N$ であるから，

$$\mathbf{z}^* A(1:k,1:k)\mathbf{z} = \mathbf{z}^* X^* A X \mathbf{z} = (X\mathbf{z})^*\, A\,(X\mathbf{z}) > 0.$$

すなわち $A(1:k,1:k)$ は正定値である．したがって系 13.30 より，$\det A(1:k,1:k) > 0$ となることがわかる．■

条件 (14.4) が成り立っていないときには次のようなことがいえる．

**定理 14.4** $A \in \boldsymbol{K}^{N\times N}$ が正則行列ならば，ある置換行列 $P$ を $PA$ が単位下三角行列 $L$ と正則な上三角行列 $U$ により $LU$ 分解可能であるようにとれる．したがって $A = P^{-1}LU$ とできる．

**証明．** まず準備的な考察から始める．$1 \leq j \leq N-1$ とし，$P = F_{j+\mu,j+\nu}$ とおく．ただし $1 \leq \mu,\nu \leq N-j$ とする．いま

$$\mathbf{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -b_{j+1} \\ \vdots \\ -b_N \end{pmatrix}$$

とし，$B = I_N - \mathbf{b}\,\mathbf{e}_j^T$ とおく．このとき

$$B = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & b_{j+1} & 1 & & & \\ O & & & 0 & \ddots & & \\ & & \vdots & \vdots & \ddots & \ddots & \\ & & b_N & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{14.6}$$

であり，したがって

$$B^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & -b_{j+1} & 1 & & & \\ O & & & 0 & \ddots & & \\ & & \vdots & \vdots & \ddots & \ddots & \\ & & -b_N & 0 & \cdots & 0 & 1 \end{pmatrix}$$

となっている．

$$PBP = P^2 - P\,\mathbf{b}\,\mathbf{e}_j^T P$$

であるが，$P^2 = I_N$ であり，$\mathbf{e}_j^T P$ は $\mathbf{e}_j^T$ の第 $j+\mu$ 列と第 $j+\nu$ 列の入れ替えなので $\mathbf{e}_j^T P = \mathbf{e}_j^T$ である．ゆえに

$$PBP = I_N - P\,\mathbf{b}\,\mathbf{e}_j^T$$

であるが，ここで $P\mathbf{b}$ の第 1 成分から第 $j$ 成分は 0 であるから，$PBP$ は (14.6) と同じ形をした行列になっている．すなわち $PBP$ の対角成分が 1 であり，それと $(j+1, j), \cdots, (N, j)$ 成分以外の成分はすべて 0 になっている．

さて，$P_1, \cdots, P_{N-1}$ を (14.2) で定めた行列とする．ある $P_k$ が単位行列ではないとしよう．このとき，以上で行った考察から，$P_k M_j P_k$ は $j < k$ のとき (14.6) と同じ形の行列になっている．これを $M_j'$ とおくと

$$\begin{aligned} P_k M_{k-1} \cdots M_1 &= P_k M_{k-1} P_k^2 M_{k-2} P_k^2 \cdots P_k^2 M_1 P_k^2 \\ &= M_{k-1}' M_{k-2}' \cdots M_1' P_k \end{aligned}$$

が成り立つ．したがって，(14.6) の形をしたある行列 $L_j$ $(1 \leq j \leq N-2)$ に対

して

$$U = M_{N-1}P_{N-1}\cdots M_1P_1A = M_{N-1}L_{N-2}\cdots L_1P_{N-1}\cdots P_1A.$$

ゆえに $P = P_{N-1}\cdots P_1$ とし，$L = L_1^{-1}\cdots L_{N-2}^{-1}M_{N-1}^{-1}$ とすると，

$$A = P^{-1}LU$$

である．ここで $P$ は置換行列であり，$L$ は単位下三角行列である．■

**注意 14.1** $A \in \boldsymbol{K}^{N\times N}$ は正則でなくとも，$LU$ 分解の過程で $a_{jj}^{(j)\prime} \neq 0,\ j = 1,\cdots,N-1$ であれば分解の操作を実行することができる．この場合，$A$ は単位下三角行列 $L$ と上三角行列 $U$ により $A = LU$ と分解できる．したがって，特に

$$\det A\,(1:k, 1:k) \neq 0, \quad k = 1,\cdots,N-1 \tag{14.7}$$

ならば，$A$ は単位下三角行列 $L$ と上三角行列 $U$ により $LU$ 分解可能である．ただしこの場合 $U$ は正則とは限らない．

**問題 14.1** $a \neq 0,\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とする．このとき次を示せ．

$$A = \begin{pmatrix} 1 & 0 \\ \dfrac{1}{a}c & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & \dfrac{1}{a}\det A \end{pmatrix}.$$

## 14.2 *LDM** 分解

$LU$ 分解をさらに分解したものが次の $LDM^*$ 分解である．

**定義 14.2** $A \in \boldsymbol{K}^{N\times N}$ がある単位下三角行列 $L,\ M \in \boldsymbol{K}^{N\times N}$ と対角行列 $D \in \boldsymbol{K}^{N\times N}$ により

$$A = LDM^* \tag{14.8}$$

と表せたとき，$A$ は $LDM^*$ **分解できる**といい，(14.8) を $A$ の $LDM^*$ 分解という．$A$ に対してこのような $L, D, M$ が一意的に存在するとき，$A$ は**一意的に** $LDM^*$ **分解できる**という．

まず次の定理を示す．

**定理 14.5** $A \in \boldsymbol{K}^{N\times N}$ が正則で $LU$ 分解可能であるとする．このとき $A$ は一

意的に $LDM^*$ 分解できる．さらに $U=(u_{ij})$ とおくと $D=\mathrm{diag}(u_{11},\cdots,u_{NN})$ で与えられる．

**証明.** 仮定より $A=LU$ となる単位下三角行列 $L$ と正則な上三角行列 $U$ が存在する．$U=(u_{jk})$ とし，$D=\mathrm{diag}(u_{11},\cdots,u_{NN})$ とおく．$u_{kk}\neq 0\,(k=1,\cdots,N)$ より $D$ は正則である．

$$M^*=D^{-1}U$$

により $M$ を定義すると，$M$ は単位下三角行列で，$A=LU=LDM^*$ を得る．$A=L'CK^*$ を別の $LDM^*$ 分解とする．$CK^*$ は正則な下三角行列であるから，$LU$ 分解の一意性から，$L=L'$ かつ $DM^*=CK^*$ であるが，$M,K$ の対角成分は 1 であるから，$D=C$ である．ゆえに $M=K$ も得られる．■

$A$ が正則なエルミート行列の場合の $LDM^*$ 分解は次のようになる．

**定理 14.6**　$A\in \boldsymbol{K}^{N\times N}$ が正則なエルミート行列で，$LU$ 分解できるとする．このとき，ある単位下三角行列 $L$ と対角行列 $D$ で

$$A=LDL^*$$

となるものが一意的に存在する．さらに $U=(u_{ij})$ とおくと $D=\mathrm{diag}(u_{11},\cdots,u_{NN})$ で与えられる．

**証明.** $LDM^*=A$ とすると，$MD^*L^*=A^*=A$ より，$M^{-1}=D^*L^*A^{-1}$ である．ゆえに

$$M^{-1}LD=D^*L^*A^{-1}LD$$

であり，$(A^{-1})^*=(A^*)^{-1}=A^{-1}$ より $M^{-1}LD$ はエルミート行列である．一方，定理 8.4 より $M^{-1}$ は単位下三角行列であるから $M^{-1}LD$ は下三角行列である．ゆえに $M^{-1}LD$ は対角行列でなければならない．したがって $M^{-1}L$ は対角行列となる．ところで，$M^{-1}$ は単位下三角行列であるから，$M^{-1}L$ の対角成分は 1 であることがわかる．よって $M=L$ が得られる．■

**注意 14.2**　対角化と異なり，$LDM^*$ 分解では，$D$ の対角成分に固有値がくるとは限らない．たとえば，$A=\begin{pmatrix}1 & 2\\ 3 & 1\end{pmatrix}$ とおくと，

$$A = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

が $LDM^*$ 分解であるが，$A$ の固有値は，$1-\sqrt{6}, 1+\sqrt{6}$ である．正則エルミートの場合も，たとえば，$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ とおくと，

$$A = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -7 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

が $LDL^*$ 分解であるが，$A$ の固有値は $\dfrac{3}{2} - \dfrac{1}{2}\sqrt{37}, \dfrac{3}{2} + \dfrac{1}{2}\sqrt{37}$ である．

ブロック行列に対して $LDM^*$ 分解に類似の分解をすることが可能である．$A \in \boldsymbol{K}^{N \times N}$ を

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

ただし, $A_{11} \in \boldsymbol{K}^{r \times r}$, $A_{22} \in \boldsymbol{K}^{(N-r)\times(N-r)}$, $A_{12} \in \boldsymbol{K}^{r\times(N-r)}$, $A_{21} \in \boldsymbol{K}^{(N-r)\times}$ とブロック分けする．$A$ および $A_{11}$ が正則であると仮定する．

$B = -A_{21}A_{11}^{-1}A_{12} + A_{22}$ とおく．基本変形を表す行列のアナロジーとして次のような行列を右からかけて，$A_{21}$ の部分を消去する．

$$\begin{aligned} \begin{pmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} &= \begin{pmatrix} A_{11} & A_{12} \\ O & -A_{21}A_{11}^{-1}A_{12} + A_{22} \end{pmatrix} \\ &= \begin{pmatrix} A_{11} & A_{12} \\ O & B \end{pmatrix}. \end{aligned}$$

また

$$\begin{pmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix}$$

より，

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ O & B \end{pmatrix}$$

が成り立つ．$D = \begin{pmatrix} A_{11} & O \\ O & B \end{pmatrix}$ とすると，$D^{-1} = \begin{pmatrix} A_{11}^{-1} & O \\ O & B^{-1} \end{pmatrix}$ であること

が容易に確認できる．

$$\begin{pmatrix} A_{11}^{-1} & O \\ O & B^{-1} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ O & B \end{pmatrix} = \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ O & I \end{pmatrix}$$

であるから，

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & B \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ O & I \end{pmatrix}$$

と分解できる．

この分解を用いればブロック行列の逆行列が次のように容易に求まる．

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1}$$
$$= \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ O & I \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & O \\ O & B^{-1} \end{pmatrix} \begin{pmatrix} I & O \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}.$$

## 14.3　コレスキー分解

$LDL^*$ 分解から次の分解定理を証明することができる．

**定理 14.7**　$A \in \boldsymbol{K}^{N\times N}$ をエルミート行列とする．このとき次の (1), (2), (3) は同値である．

(1)　$A$ は正定値である．

(2)　任意の $k = 1, \cdots, N$ に対して

$$\det A\,(1:k, 1:k) > 0$$

が成り立つ．

(3)　対角成分が正の下三角行列 $G$ で

$$A = GG^*$$

となるものが存在する．

なおもし (3) の分解が成り立つならば，その分解は一意的である．

**証明．** (1) ⇒ (2) は定理 14.3 で証明した．

(2) ⇒ (3) を示す．定理 14.6，仮定 (2) および定理 14.2 より，$A = LDL^*$ とで

きる．ここで $L$ は単位下三角行列，$D$ は対角行列である．定理 14.6 の後半の主張と (14.5) および仮定 (2) より $D=\mathrm{diag}(d_1,\cdots,d_N)$ とおくと，$d_k>0$ であることがわかる $(k=1,\cdots,N)$．そこで，

$$G = L\,\mathrm{diag}\left(\sqrt{d_1},\cdots,\sqrt{d_N}\right)$$

とおくと，$G$ は対角成分が $\sqrt{d_1},\cdots,\sqrt{d_N}$ の下三角行列で，

$$GG^* = LDL^* = A$$

である．

(3) $\Rightarrow$ (1) を示す．仮定 (3) の分解 $A=GG^*$ において，$G=(g_{ij})$ とすると $G$ が三角行列であることと $G$ の対角成分が正であることより

$$\det G = u_{11}\cdots u_{NN} > 0$$

を得る．したがって $G$ は正則行列である．ゆえに任意の $\mathbf{x}\in\boldsymbol{K}^N$, $\|\mathbf{x}\|\neq 0$ に対して $G^*\mathbf{x}\neq\mathbf{0}$ であり，

$$\mathbf{x}^*A\mathbf{x} = \mathbf{x}^*GG^*\mathbf{x} = (G^*\mathbf{x})^*\,(G^*\mathbf{x}) = \|G^*\mathbf{x}\|^2 > 0$$

が成り立つ．ゆえに $A$ は正定値である．

最後に (3) の分解の一意性を証明する．$A=KK^*$ を別の同様な分解とする．このとき

$$\begin{aligned} G^{-1}K &= G^{-1}\left(AA^{-1}\right)K = G^{-1}GG^*A^{-1}K \\ &= G^*A^{-1}K = G^*\left(K^{-1}\right)^*. \end{aligned}$$

ここで $G^{-1}K$ は下三角行列，$G^*\left(K^{-1}\right)^*$ は上三角行列であるから，$G^{-1}K=\Delta$ は対角行列で，その対角成分は正である．いま

$$GG^* = A = KK^* = G\Delta\Delta^*G^* = G\Delta\Delta G^*.$$

したがって，$\Delta\Delta=I_N$ で，$\Delta$ の対角成分は正であるから，$\Delta=I_N$ となる．よって $G=K$．■

**定義 14.3**　定理 14.7 (3) における $A$ の分解 $A=GG^*$ を $A$ のコレスキー分解といい，$G$ を $A$ の**コレスキー下三角行列**という．

## 14.4　$QR$ 分解

本節ではまず行列 $A \in \boldsymbol{K}^{M\times N}$ をユニタリー行列 $Q \in \boldsymbol{K}^{M\times M}$ と上三角行列 $R \in \boldsymbol{K}^{M\times N}$ により $A = QR$ と分解できることを示す．次に少し異なったタイプの分解を示す．

### 14.4.1　ハウスホルダー $QR$ 分解

ベクトル $\mathbf{v} \in \boldsymbol{K}^M$, $\mathbf{v} \neq \mathbf{0}$ に関するハウスホルダー行列 (例 2.5 参照) を

$$H_{\mathbf{v}} = I_M - \frac{2}{\|\mathbf{v}\|^2}\mathbf{v}\mathbf{v}^*$$

とする．例 2.5 より次のことが示されている．

**命題 14.8**　ハウスホルダー行列はエルミートかつユニタリー行列である．

まず，ベクトル $\mathbf{x} \in \boldsymbol{K}^M$ をあるハウスホルダー行列により，第 1 成分以外は 0 となるベクトル $\pm\|\mathbf{x}\|\,\mathbf{e}_1^{(M)}$ に変換できることを示す．

**定理 14.9**　$\mathbf{x} = (x_j) \in \boldsymbol{K}^M$, $\mathbf{x} \neq \mathbf{0}$ とする．$j = 1, \cdots, M$ に対して，$\mathbf{v} = \mathbf{x} + \varepsilon\|\mathbf{x}\|\,\mathbf{e}_j^{(M)}$ とする．ただし，ここで $x_j = 0$ ならば $\varepsilon = 1$, $x_j \neq 0$ ならば $\varepsilon = \dfrac{x_j}{|x_j|}$ とおく．このとき，

$$H_{\mathbf{v}}\mathbf{x} = -\varepsilon\|\mathbf{x}\|\,\mathbf{e}_j^{(M)}.$$

**証明.**

$$\|\mathbf{v}\|^2 = \left(\mathbf{x} + \varepsilon\|\mathbf{x}\|\,\mathbf{e}_j^{(M)}\right)^*\left(\mathbf{x} + \varepsilon\|\mathbf{x}\|\,\mathbf{e}_j^{(M)}\right) = 2\left(\|\mathbf{x}\|^2 + \varepsilon\|\mathbf{x}\|\,\overline{x_j}\right),$$
$$\mathbf{v}^*\mathbf{x} = \|\mathbf{x}\|^2 + \varepsilon\|\mathbf{x}\|\,\overline{x_j}$$

である．ゆえに

$$H_{\mathbf{v}}\mathbf{x} = \mathbf{x} - \frac{2\left(\|\mathbf{x}\|^2 + \varepsilon\|\mathbf{x}\|\,\overline{x_j}\right)}{2\left(\|\mathbf{x}\|^2 + \varepsilon\|\mathbf{x}\|\,\overline{x_j}\right)}\mathbf{v} = -\varepsilon\|\mathbf{x}\|\,\mathbf{e}_j^{(M)}$$

が得られる．■

このことを用いて，いわゆる $QR$ 分解と呼ばれる行列の分解が得られる．

**定理 14.10**　$A \in \boldsymbol{K}^{M\times N}$ とし，$L = \min\{M, N\}$ とする．あるユニタリー行列 $Q \in \boldsymbol{K}^{M\times M}$ と上三角行列 $R = (r_{jk}) \in \boldsymbol{K}^{M\times N}$ で，$r_{jj} \geq 0\,(j = 1, \cdots, L)$ をみたし，かつ

$$A = QR$$

となるものが存在する．

**証明．** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．$\mathbf{a}_1 = \mathbf{0}$ ならば $P_1 = I_M$ とおく．$\mathbf{a}_1 \neq \mathbf{0}$ ならば，$P_1$ として $\mathbf{a}_1$ を定理 14.9 のように $-\varepsilon_1 \|\mathbf{a}_1\| \mathbf{e}_1^{(M)}$ に写すハウスホルダー行列を考える．これにより

$$P_1 A = \left(\mathbf{a}_1^{(1)}, \cdots, \mathbf{a}_N^{(1)}\right) = \begin{pmatrix} -\varepsilon_1 \|\mathbf{a}_1\| & & & \\ 0 & \mathbf{a}_2^{(1)} & \cdots & \mathbf{a}_N^{(1)} \\ \vdots & & & \\ 0 & & & \end{pmatrix}$$

とできる．$\widetilde{\mathbf{a}}_2^{(2)} = \mathbf{a}_2^{(1)}(2:M)$ とする．もし $\widetilde{\mathbf{a}}_2^{(2)} = \mathbf{0}_{M-1}$ ならば $P_2 = I_M$ とする．また $\widetilde{\mathbf{a}}_2^{(2)} \neq \mathbf{0}_{M-1}$ ならば，$\widetilde{P}_2 \in \boldsymbol{K}^{(M-1)\times(M-1)}$ として $\widetilde{\mathbf{a}}_2^{(2)}$ を定理 14.9 のように $-\varepsilon_2 \left\|\widetilde{\mathbf{a}}_2^{(2)}\right\| \mathbf{e}_1^{(M-1)}$ に写すハウスホルダー変換を考える．$P_2 = \begin{pmatrix} 1 & O \\ O & \widetilde{P}_2 \end{pmatrix}$ とおくと，

$$P_2 P_1 A = \begin{pmatrix} -\varepsilon_1 \|\mathbf{a}_1\| & * & & & \\ 0 & -\varepsilon_2 \left\|\widetilde{\mathbf{a}}_2^{(2)}\right\| & & & \\ 0 & 0 & \mathbf{a}_3^{(2)} & \cdots & \mathbf{a}_N^{(2)} \\ \vdots & \vdots & & & \\ 0 & 0 & & & \end{pmatrix}$$

とできる．ただし $*$ はあるスカラーを表す．$\widetilde{\mathbf{a}}_3^{(3)} = \mathbf{a}_3^{(2)}(3:M)$ とする．$\widetilde{\mathbf{a}}_3^{(3)} = \mathbf{0}_{M-2}$ ならば $P_3 = I_M$ とする．また $\widetilde{\mathbf{a}}_3^{(3)} \neq \mathbf{0}_{M-2}$ ならば，$\widetilde{P}_3 \in \boldsymbol{K}^{(M-2)\times(M-2)}$ として $\widetilde{\mathbf{a}}_3^{(3)}$ を定理 14.9 のように $-\varepsilon_3 \left\|\widetilde{\mathbf{a}}_3^{(3)}\right\| \mathbf{e}_1^{(M-2)}$ に写すハウスホルダー行列を考える．$P_2 = \begin{pmatrix} I_2 & O \\ O & \widetilde{P}_3 \end{pmatrix}$ とおくと，

$$P_3 P_2 P_1 A = \begin{pmatrix} -\varepsilon_1 \|\mathbf{a}_1\| & * & * & & & \\ 0 & -\varepsilon_2 \left\|\widetilde{\mathbf{a}}_2^{(2)}\right\| & * & & & \\ 0 & 0 & -\varepsilon_3 \left\|\widetilde{\mathbf{a}}_3^{(3)}\right\| & & & \\ 0 & 0 & 0 & \mathbf{a}_4^{(3)} & \cdots & \mathbf{a}_N^{(3)} \\ \vdots & \vdots & \vdots & & & \\ 0 & 0 & 0 & & & \end{pmatrix}$$

となる．この操作を繰り返せば $P_L P_{L-1}\cdots P_1 A$ を上三角行列にできる．

ここで，さらに $D = \mathrm{diag}_M\left(-\overline{\varepsilon_1},\cdots,-\overline{\varepsilon_L}\right)$ とおけば，$DP_L\cdots P_1 A$ は上三角行列でその対角成分は $\|\mathbf{a}_1\|$, $\left\|\widetilde{\mathbf{a}}_2^{(2)}\right\|$, $\cdots$, $\left\|\widetilde{\mathbf{a}}_L^{(L)}\right\|$ である．$Q = P_1^{-1}\cdots P_L^{-1}D^{-1}$, $R = DP_L\cdots P_1 A$ が求める分解である．■

この定理の変形として，ユニタリー行列と帯行列と呼ばれる行列への分解も成り立つ．定理を述べるために帯行列の定義を述べる．

$(M,N)$ 行列 $A = (a_{jk})$ が**下帯幅** $p$ **の帯行列**とは，$a_{jk} = 0$, $j-k > p$ をみたすことである．また**上帯幅** $q$ **の帯行列**とは $a_{jk} = 0$, $k-j > q$ をみたすことである．特に $a_{jk} = 0$, $j-k > p$, $k-j > q$ をみたすとき**下帯幅** $p$ **かつ上帯幅** $q$ **の帯行列**という．たとえば

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & 0 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} \end{pmatrix}, \quad B = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{pmatrix}$$

とすると，$A$ は下帯幅 2 かつ上帯幅 1 の帯行列である．また $B$ は上帯幅 1 の帯行列である．

$(M,N)$ 行列が上三角行列ならば下帯幅 0, 上帯幅 $N-1$ の帯行列であり，下三角行列は下帯幅 $M-1$，上帯幅 0 の帯行列である．

なお，帯幅が負数の場合も考える．たとえば，

$$C = \begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \\ a_{41} & a_{42} & a_{43} \end{pmatrix}$$

は上帯幅 $-1$ の帯行列である．

上帯幅 $N-M$ の帯行列を本書では**後ろ向き下三角行列**という．

(14.9) の行列のように後ろ向き下三角行列で，さらに $a_{jk} = 0$, $k-j < N-M$ もみたすものを**後ろ向き対角行列**と呼び，$a_{jk}$, $k-j = N-M$ を**後ろ向き対角成分**と呼ぶ．

後向き対角成分が $\beta_1,\cdots,\beta_{M\wedge N}$ の後向き対角行列を，$\mathrm{bdiag}_{M,N}\left(\beta_1,\cdots,\beta_{M\wedge N}\right)$ と表す．

$$\begin{pmatrix} 0 & \beta_1 & 0 & 0 & 0 \\ 0 & 0 & \beta_2 & 0 & 0 \\ 0 & 0 & 0 & \beta_3 & 0 \\ 0 & 0 & 0 & 0 & \beta_4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ \beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ 0 & 0 & \beta_3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \tag{14.9}$$

**定理 14.11**　$A \in \boldsymbol{K}^{M\times N}$ とし，$N \leq M$ とする．あるユニタリー行列 $Q \in \boldsymbol{K}^{M\times M}$ と上帯幅 $N-M$ の帯行列 $S = (r_{jk}) \in \boldsymbol{K}^{M\times N}$ で，$r_{jk} \geq 0$, $j-k = M-N$ をみたし，かつ

$$A = QS$$

となるものが存在する．

**証明．** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．$\mathbf{a}_N = \mathbf{0}$ ならば $P_N = I_M$ とおく．$\mathbf{a}_N \neq \mathbf{0}$ ならば，$P_N$ として $\mathbf{a}_N$ を定理 14.9 のように $-\varepsilon_N \|\mathbf{a}_N\| \mathbf{e}_M^{(M)}$ に写すハウスホルダー行列を考える．これにより

$$P_N A = \left(\mathbf{a}_1^{(1)}, \cdots, \mathbf{a}_N^{(1)}\right) = \begin{pmatrix} & & & 0 \\ \mathbf{a}_1^{(1)} & \cdots & \mathbf{a}_{N-1}^{(1)} & \vdots \\ & & & 0 \\ & & & -\varepsilon_N \|\mathbf{a}_N\| \end{pmatrix}$$

とできる．$\widetilde{\mathbf{a}}_{N-1}^{(2)} = \mathbf{a}_{N-1}^{(1)}(1:M-1)$ とする．もし $\widetilde{\mathbf{a}}_{N-1}^{(2)} = \mathbf{0}_{M-1}$ ならば $P_{N-1} = I_{M-1}$ とする．また $\widetilde{\mathbf{a}}_{N-1}^{(2)} \neq \mathbf{0}_{M-1}$ ならば, $\widetilde{P}_{N-1} \in \boldsymbol{K}^{(M-1)\times(M-1)}$ により $\widetilde{\mathbf{a}}_{N-1}^{(2)}$ を定理 14.9 のように $-\varepsilon_{N-1} \left\|\widetilde{\mathbf{a}}_{N-1}^{(2)}\right\| \mathbf{e}_{M-1}^{(M-1)}$ に写すハウスホルダー行列を考える．$P_{N-1} = \begin{pmatrix} \widetilde{P}_{N-1} & O \\ O & 1 \end{pmatrix}$ とおくと，

$$P_{N-1}P_N A = \begin{pmatrix} & & 0 & 0 \\ & & \vdots & \vdots \\ \mathbf{a}_1^{(2)} & \mathbf{a}_{N-2}^{(2)} & 0 & 0 \\ & & -\varepsilon_{N-1} \left\|\widetilde{\mathbf{a}}_{N-1}^{(2)}\right\| & 0 \\ & & * & -\varepsilon_N \|\mathbf{a}_N\| \end{pmatrix}$$

とできる．$\widetilde{\mathbf{a}}_{N-2}^{(3)} = \mathbf{a}_{N-2}^{(2)}(1:M-2)$ とする．$\widetilde{\mathbf{a}}_{N-2}^{(3)} = \mathbf{0}_{M-2}$ ならば $P_{N-2} = I_{M-2}$ とする．また $\widetilde{\mathbf{a}}_{N-2}^{(3)} \neq \mathbf{0}_{M-2}$ ならば，$\widetilde{P}_{N-2} \in \boldsymbol{K}^{(M-2)\times(M-2)}$ により $\widetilde{\mathbf{a}}_{N-2}^{(3)}$ を定理 14.9 のように $-\varepsilon_{N-2} \left\|\widetilde{\mathbf{a}}_{N-2}^{(3)}\right\| \mathbf{e}_{M-2}^{(M-2)}$ に写すハウスホルダー行列を

考える．$P_{N-2} = \begin{pmatrix} \widetilde{P}_{N-2} & O \\ O & I_2 \end{pmatrix}$ とおくと，

$$P_{N-2}P_{N-1}P_N A$$
$$= \begin{pmatrix} & & 0 & 0 & 0 \\ & & \vdots & \vdots & \vdots \\ & & 0 & 0 & 0 \\ \mathbf{a}_1^{(3)} & \mathbf{a}_{N-3}^{(3)} & -\varepsilon_{N-2}\left\|\widetilde{\mathbf{a}}_{N-2}^{(3)}\right\| & 0 & 0 \\ & & * & -\varepsilon_{N-1}\left\|\widetilde{\mathbf{a}}_{N-1}^{(2)}\right\| & 0 \\ & & & * & -\varepsilon_N \left\|\mathbf{a}_N\right\| \end{pmatrix}$$

となる．この操作を繰り返せば $P_1 \cdots P_N A$ を上帯幅 $N-M$ の帯行列にできる．ここでさらに $D = \mathrm{diag}_M\,(1, \cdots, 1, -\overline{\varepsilon_1}, \cdots, -\overline{\varepsilon_N})$ とおけば，$DP_1 \cdots P_N A$ は上帯幅 $N-M$ の帯行列でその $(j, j+(M-N))$ 成分は $\left\|\widetilde{\mathbf{a}}_1^{(N)}\right\|, \cdots, \left\|\widetilde{\mathbf{a}}_{N-1}^{(2)}\right\|, \|\mathbf{a}_N\|$ である．$Q = P_N^{-1} \cdots P_1^{-1} D^{-1}$, $S = DP_1 \cdots P_N A$ が求める分解である．■

**問題 14.2**　$N > M$ の場合について考察せよ．

### 14.4.2　シュミットの方法による *QR* 分解

シュミットの直交化を用いて次の分解を得ることができる．

**定理 14.12**　$N \leq M$, $A \in \boldsymbol{K}^{M \times N}$, $\mathrm{rank}(A) = N$ とする．このときある正規直交系 $\mathbf{u}_1, \cdots, \mathbf{u}_N \in \boldsymbol{K}^M$ とある上三角行列 $R \in \boldsymbol{K}^{N \times N}$ で，対角成分が 1 からなり，かつ

$$A = QR, \quad \text{ただし} \quad Q = (\mathbf{u}_1, \cdots, \mathbf{u}_N)$$

をみたすものがとれる．

**証明.** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とすると，$\mathbf{a}_1, \cdots, \mathbf{a}_N$ は線形独立である．このとき，シュミットの直交化法を用いて

$$\begin{aligned} \mathbf{u}_1 &= r_{11}\mathbf{a}_1, \\ \mathbf{u}_2 &= r_{21}\mathbf{a}_1 + r_{22}\mathbf{a}_2, \\ &\vdots \\ \mathbf{u}_N &= r_{N1}\mathbf{a}_1 + \cdots + r_{NN}\mathbf{a}_N \end{aligned}$$

の係数 $r_{jk}$ を定めることにより，正規直交系 $\mathbf{u}_1, \cdots, \mathbf{u}_N$ を作ることができる．ただし，ここで $r_{kk} = 1$ となるようにできる．したがって，

$$R = \begin{pmatrix} r_{11} & r_{21} & \cdots & r_{N1} \\ 0 & r_{22} & \cdots & r_{N2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{NN} \end{pmatrix}^{-1}$$

とすると，$A = QR$ である． ■

# 第3部
# 応用編

# 第15章　一般逆行列とその応用

この章では，逆行列の概念を一般化した一般逆行列について述べる．一般逆行列は正則でない行列，あるいは正方行列でない行列の解析で威力を発揮する．また一般逆行列の応用例として，解を持たない連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ に対して，$\|A\mathbf{x} - \mathbf{b}\|$ を最小にする $\mathbf{x}$ を求めることや，与えられたデータを最もよく近似する直線を求めることができる．

**ポイント 15**　一般逆行列は多くの線形代数の教科書では取り上げられることが少ないが，有用な考え方であると同時に応用性にも富んでいる．

## 15.1　一般逆行列

$N$ 次正方行列 $A$ が正則であるとは，逆行列 $A^{-1}$ が存在することである．$A$ が正則でない場合，あるいは正方行列でない場合，逆行列は存在しない．逆行列の概念を一般化し，正則あるいは正方でない行列に対しても逆行列に類似の概念を導入する．

**定義 15.1**　$A \in \boldsymbol{K}^{M \times N}$ とする．$A^- \in \boldsymbol{K}^{N \times M}$ で，

$$AA^-A = A \tag{15.1}$$

をみたすものを $A$ の**一般逆行列**という．

もし $M = N$ で，$A$ が正則であれば，$A$ の逆行列 $A^{-1}$ は

$$AA^{-1}A = AI_M = A$$

をみたし，$A$ の一般逆行列になっている．逆に $A$ が正則である場合，$A$ の一般逆行列 $A^-$ は

$$AA^{-} = AA^{-}\left(AA^{-1}\right) = \left(AA^{-}A\right)A^{-1} = AA^{-1} = I_N$$

をみたしている．したがって $A^{-}$ は $A$ の逆行列である．このことから正則行列に対しては一般逆行列と逆行列は同じものとなっていることがわかる．したがって一般逆行列の研究は，$A$ が正則でない場合，あるいは $M \neq N$ の場合に行われる．

逆行列と異なり，一般逆行列はどのような行列に対しても存在する．

**定理 15.1**　$A \in \boldsymbol{K}^{M\times N}$ とする．このとき，$A$ の一般行列 $A^{-}$ が存在する．

**証明.** $r = \mathrm{rank}(A)$ とする．定理 11.17 よりある $M$ 次正則行列 $E$，$N$ 次置換行列 $P$ が存在し，

$$EAP = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix}$$

となる ( ただし $K$ はある $(r, N-r)$ 行列．規約 11.16 参照)．$S = EAP$ とおく．さて，$L$ を任意の $(N-r, M-r)$ 行列とし，

$$\Sigma = \begin{pmatrix} I_r & O \\ O & L \end{pmatrix}$$

とする．ただし，$M-r=0$ または $N-r=0$ の場合は規約 11.16 に従うものとする．このとき

$$\begin{aligned} S\Sigma S &= \begin{pmatrix} I_r & K \\ O & O \end{pmatrix}\begin{pmatrix} I_r & O \\ O & L \end{pmatrix}\begin{pmatrix} I_r & K \\ O & O \end{pmatrix} \\ &= \begin{pmatrix} I_r & KL \\ O & O \end{pmatrix}\begin{pmatrix} I_r & K \\ O & O \end{pmatrix} = \begin{pmatrix} I_r & K \\ O & O \end{pmatrix} = S \end{aligned}$$

であるから，$X = P\Sigma E$ とおけば

$$\begin{aligned} AXA &= \left(E^{-1}SP^{-1}\right)(P\Sigma E)\left(E^{-1}SP^{-1}\right) = E^{-1}S\Sigma SP^{-1} \\ &= E^{-1}SP^{-1} = A \end{aligned}$$

が成り立つ．よって $X$ は $A$ の一般逆行列である．■

$A$ の一般逆行列 $A^{-}$ は存在するが，必ずしも一意的とは限らない．このことは証明中で $\Sigma$ を作るときに $L$ に関する任意性があったことからわかるであろう (たとえば次の例題 15.1 の答を見よ)．

**例題 15.1**　定理 15.1 の証明を参考にして，$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$ の一般逆行列を求めよ．

**答．** $A$ は行に関する基本変形により，

$$A \xrightarrow{E_{3,1}(-1)} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \xrightarrow{E_{3,2}(-1)} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{E_{1,2}(-1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

と階段行列に変形できる．ただし記号については第 4.3.2 節を参照．また $A \xrightarrow{B} C$ は $BA$ が $C$ になることを意味している．したがって定理 15.1 の証明中の $E$ は

$$E = E_{1,2}(-1)\,E_{3,2}(-1)\,E_{3,1}(-1) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

である．また $P = I_3$ であり，$\Sigma$ は任意の複素数 $a$ に対して

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix}$$

として得られる．これより

$$A^- = I_3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -a & -a & a \end{pmatrix}$$

である．■

**例題 15.2**　$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \end{pmatrix}$ の一般逆行列を一つ求めよ．

**答．** $A$ は行に関する基本変形により，

$$A \xrightarrow{E_{2,1}(-2)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \end{pmatrix} \xrightarrow{E_{2}(-1)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\xrightarrow{E_{1,2}(-2)} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \end{pmatrix}$$

のようにして標準形にできる．したがって

$$E = E_{1,2}(-2)\, E_2(-1) E_{2,1}(-2) = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix},$$

$$P = I_3$$

とすると，

$$A^- = I_3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix}$$

である． ■

ある一つの一般逆行列がわかれば，それから他のすべての一般逆行列を次のようにして求めることができる．

**定理 15.2**　$A \in \boldsymbol{K}^{M\times N}$ とする．$A$ の一般逆行列 $A^-$ を任意にとり固定する．このとき，任意の $B \in \boldsymbol{K}^{N\times M}$ に対して，

$$A^- + B - A^- ABAA^- \tag{15.2}$$

は $A$ の一般逆行列になっている．逆に，$A$ の任意の一般逆行列はある $B \in \boldsymbol{K}^{N\times M}$ により (15.2) と表せる．

**証明.** まず前半の主張を証明する．任意の $B \in \boldsymbol{K}^{N\times M}$ に対して $A' = A^- + B - A^- ABAA^-$ とおくと，

$$AA'A = A\left(A^- + B - A^- ABAA^-\right) A = A + ABA - ABA = A$$

である．次に後半の主張を証明する．$A'$ を $A$ の一般逆行列とする．$B = A' - A^-$ とおくと，

$$ABA = AA'A - AA^- A = A - A = O$$

であるから，$B - A^- ABAA^- = B$．ゆえに

$$A' = A^- + B = A^- + B - A^- ABAA^-$$

と表せている． ■

一般逆行列を用いて，可解な非斉次方程式 $A\mathbf{x} = \mathbf{b}$ の解の一般的な形を記述することができる．

**定理 15.3**　$A \in \boldsymbol{K}^{M\times N}$, $\mathbf{b} \in \boldsymbol{K}^{M}$ とする．このとき，方程式

$$A\mathbf{x} = \mathbf{b} \tag{15.3}$$

が可解ならば，$A$ のすべての一般逆行列 $A^-$ に対して，

$$AA^-\mathbf{b} = \mathbf{b} \tag{15.4}$$

が成り立つ．また $A$ の任意の一般逆行列 $A^-$ と任意の $\mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x} = A^-\mathbf{b} + (I - A^-A)\mathbf{y} \tag{15.5}$$

が (15.3) の解になっている．逆に (15.3) の解は (15.5) の形で表すことができる．

**証明.** (15.3) が可解ならば，$\mathbf{x}$ を $A\mathbf{x} = \mathbf{b}$ の解とするとき，

$$AA^-\mathbf{b} = AA^-A\mathbf{x} = A\mathbf{x} = \mathbf{b}$$

が成り立つ．またこのとき，任意の $\mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\begin{aligned} A\left(A^-\mathbf{b} + (I - A^-A)\mathbf{y}\right) &= \mathbf{b} + A\mathbf{y} - AA^-A\mathbf{y} \\ &= \mathbf{b} + A\mathbf{y} - A\mathbf{y} = \mathbf{b} \end{aligned}$$

が成り立っている．一方，$\mathbf{x}$ を (15.3) の解とすると，$A^-\mathbf{b} = A^-A\mathbf{x}$ であるから

$$\mathbf{x} = A^-\mathbf{b} + \mathbf{x} - A^-A\mathbf{x} = A^-\mathbf{b} + (I - A^-A)\mathbf{x}$$

と表せている．■

**問題 15.1**　$A \in \boldsymbol{K}^{M\times N}$, $X \in \boldsymbol{K}^{N\times M}$ とする．次の (1), (2) が同値であることを証明せよ．

(1)　$X$ が $A$ の一般逆行列．

(2)　$A\mathbf{x} = \mathbf{b}$ が可解であるような任意の $\mathbf{b} \in \boldsymbol{K}^N$ に対して，$AX\mathbf{b} = \mathbf{b}$.

$A \in \boldsymbol{K}^{M\times N}$ に対して，$X \in \boldsymbol{K}^{N\times M}$ で

$$AX = I_M$$

をみたすものが存在するとき，$X$ を $A$ の**右逆行列**という．また $Y \in \boldsymbol{K}^{N\times M}$ で

$$YA = I_N$$

をみたすものが存在するとき，$Y$ を $A$ の**左逆行列**という．右あるいは左逆行列は

明らかに一般逆行列になっている．

ところで，すでに示したように，$M = N$ の場合，$A$ の右逆行列あるいは左逆行列が存在するならば，$A$ は正則で，右逆行列，左逆行列ともに存在し，それらはすべて $A$ の逆行列と一致する．しかし $M \neq N$ の場合は，次のことが成り立つ．

**命題 15.4** $A \in \boldsymbol{K}^{M \times N}$ で，$M \neq N$ とする．もし右逆行列が存在するならば $\mathrm{rank}(A) = M$ であり，左逆行列が存在するならば $\mathrm{rank}(A) = N$ である．したがって，もし右逆行列が存在するならば左逆行列は存在しない．また左逆行列が存在するならば右逆行列は存在しない．

**証明.** もし右逆行列 $X$ が存在するならば，

$$M = \mathrm{rank}(I_M) = \mathrm{rank}(AX) \leq \mathrm{rank}(A) \leq \min\{M, N\} \leq M$$

より $\mathrm{rank}(A) = M$ である．同様に，もし左逆行列が存在するならば $\mathrm{rank}(A) = N$ である．これより命題が証明される． ■

以下に示すように，右あるいは左逆行列が存在する場合，すべての一般逆行列は右あるいは左逆行列になっている．

**定理 15.5** $A \in \boldsymbol{K}^{M \times N}$ とする．次のことが成り立つ．

(1) $A$ の右逆行列が存在するための必要十分条件は $\mathrm{rank}(A) = M$ となることである．さらにこのとき，$A$ の一般逆行列は $A$ の右逆行列になっている．

(2) $A$ の左逆行列が存在するための必要十分条件は $\mathrm{rank}(A) = N$ となることである．さらにこのとき，$A$ の一般逆行列は $A$ の左逆行列になっている．

証明をする前に次の補題を示す．

**補題 15.6** $A \in \boldsymbol{K}^{M \times N}$ とし，$A^-$ を一般逆行列とする．このとき，

$$\left(AA^-\right)\left(AA^-\right) = AA^-, \quad \left(A^-A\right)\left(A^-A\right) = A^-A, \tag{15.6}$$

$$\mathrm{rank}(AA^-) = \mathrm{rank}(A^-A) = \mathrm{rank}(A) \tag{15.7}$$

が成り立っている．

**証明.**

$$\left(AA^-\right)\left(AA^-\right) = \left(AA^-A\right)A^- = AA^-,$$
$$\left(A^-A\right)\left(A^-A\right) = A^-\left(AA^-A\right) = A^-A$$

である．また

$$\mathrm{rank}(A) = \mathrm{rank}(AA^-A) \leq \mathrm{rank}(A^-A) \leq \mathrm{rank}(A),$$
$$\mathrm{rank}(A) = \mathrm{rank}(AA^-A) \leq \mathrm{rank}(AA^-) \leq \mathrm{rank}(A)$$

となっている．■

**定理 15.5 の証明.** (1)　$A$ の右逆行列が存在するならば，命題 15.4 の証明より $\mathrm{rank}(A) = M$ である．逆に，$\mathrm{rank}(A) = M$ であるとする．このとき $AA^-$ は $M$ 次正方行列で，(15.7) より $\mathrm{rank}(AA^-) = M$ であり，したがって $AA^-$ は正則である．ゆえに (15.6) より

$$AA^- = \left(AA^-\right)^{-1}\left(AA^-\right)\left(AA^-\right) = \left(AA^-\right)^{-1}\left(AA^-\right) = I_M$$

となっている．(2) も同様にして証明できる．■

## 15.2　ムーア-ペンローズ一般逆行列

本節では一般逆行列の中で，特にムーア-ペンローズ一般逆行列と呼ばれているものについて学ぶ．

**定義 15.2**　$A \in \boldsymbol{K}^{M\times N}$ とする．このとき $A^\dagger \in \boldsymbol{K}^{N\times M}$ が条件

(1)　$AA^\dagger A = A,$

(2)　$A^\dagger AA^\dagger = A^\dagger,$

(3)　$(AA^\dagger)^* = AA^\dagger,$

(4)　$(A^\dagger A)^* = A^\dagger A$

をみたすとき，$A^\dagger$ を $A$ の**ムーア-ペンローズ一般逆行列**(略して MP **一般逆行列**) という．

ムーア-ペンローズ一般逆行列の条件のうち，(1) は $A^\dagger$ が $A$ の一般逆行列であることを意味している．(2) は逆に $A$ が $A^\dagger$ の一般逆行列であることを保障する条件である．さて，いま $N = M$ で $A$ が正則であれば $AA^{-1} = A^{-1}A = I_N$ であるが，このことは $AA^{-1}$, $A^{-1}A$ がそれぞれ $\boldsymbol{K}^N$ から $\boldsymbol{K}^N\,(=\mathcal{R}(A))$ あるいは $\boldsymbol{K}^N$ から $\boldsymbol{K}^N\,(=\mathcal{R}(A^{-1}))$ への直交射影であることにほかならない．一般の場合のムーア-ペンローズ一般逆行列については，条件 (1), (2) から

$$\left(AA^\dagger\right)\left(AA^\dagger\right) = AA^\dagger,$$
$$\left(A^\dagger A\right)\left(A^\dagger A\right) = A^\dagger A$$

であるから，$AA^\dagger$ および $A^\dagger A$ は射影行列になっている．条件 (3), (4) はさらに $AA^\dagger$ および $A^\dagger A$ が直交射影行列であることを保障する条件である．また

$$\mathcal{R}(A) = \mathcal{R}\left(AA^\dagger A\right) \subset \mathcal{R}\left(AA^\dagger\right) \subset \mathcal{R}(A),$$
$$\mathcal{R}\left(A^\dagger\right) = \mathcal{R}\left(A^\dagger AA^\dagger\right) \subset \mathcal{R}\left(A^\dagger A\right) \subset \mathcal{R}\left(A^\dagger\right)$$

より，$AA^\dagger$ は $\boldsymbol{K}^M$ から $\mathcal{R}(A)$ への直交射影，そして $A^\dagger A$ は $\boldsymbol{K}^N$ から $\mathcal{R}\left(A^\dagger\right)$ への直交射影になっている．

本節では，$A$ のムーア-ペンローズ一般逆行列が存在することおよび一意的であることを証明する．そのために階数分解定理を示す．

**定理 15.7**　$A \in \boldsymbol{K}^{M\times N}$ とし，$r = \mathrm{rank}(A)$ とする．このとき，ある $B \in \boldsymbol{K}^{M\times r}$ とある $C \in \boldsymbol{K}^{r\times N}$ で，

$$A = BC,$$
$$\mathrm{rank}(B) = \mathrm{rank}(C) = r$$

をみたすものをとることができる．$A = BC$ を $A$ の**階数分解**という．

**証明.** $A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．$\mathrm{rank}(A) = r$ より，$r$ 個の線形独立なベクトル $\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}$ をとることができる．置換行列 $P$ をとり，

$$AP = \left(\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}, \mathbf{a}'_{r+1}, \cdots, \mathbf{a}'_N\right)$$

とできる．$A_1 = \left(\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}\right)$, $A_2 = \left(\mathbf{a}'_{r+1}, \cdots, \mathbf{a}'_N\right)$ とおく．$\mathrm{rank}(A) = r$ であるから，$\mathbf{a}'_j \in \mathrm{span}\left\{\mathbf{a}_{i(1)}, \cdots, \mathbf{a}_{i(r)}\right\}$, $j = r+1, \cdots, N$ であり，したがって，ある $Q \in \boldsymbol{K}^{r\times(N-r)}$ を

$$A_2 = A_1 Q$$

となるようにとれる．

$$B = A_1, \quad C = (I_r, Q)\, P^{-1}$$

とおくと，$\mathrm{rank}(B) = \mathrm{rank}(C) = r$ であり，

$$BC = A_1 (I_r, Q)\, P^{-1} = (A_1, A_2)\, P^{-1} = A$$

となっている．■

**定理 15.8**　$A \in \boldsymbol{K}^{M\times N}$ とする．このとき $A$ のムーア-ペンローズ一般逆行列 $A^\dagger$ が一意的に存在する．特に $A = BC$ を定理 15.7 の分解とすると，

$$A^{\dagger} = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$$

で与えられる．

**証明．** まず存在を証明する．$A = BC$ を定理 15.7 の分解とする．このとき系 11.13 より，$B^*B$ は正則である．同様にして $CC^*$ も正則である．したがって，$B' = (B^*B)^{-1}B^*$, $C' = C^*(CC^*)^{-1}$ とすると，

$$B'B = I_r, \quad CC' = I_r$$

より，$B', C'$ はそれぞれ $B, C$ の左逆行列，右逆行列になっている．

$$A^{\dagger} = C'B'$$

とおくと，

$$\begin{aligned} AA^{\dagger}A &= BCC'B'BC = BC = A, \\ A^{\dagger}AA^{\dagger} &= C'B'BCC'B' = C'B' = A^{\dagger} \end{aligned}$$

である．さらに

$$\begin{aligned} AA^{\dagger} &= BCC'B' = BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* \\ &= B(B^*B)^{-1}B^* \end{aligned}$$

であるから，$(B^*B)^* = B^*B$ より

$$\left(AA^{\dagger}\right)^* = B\,(B^*B)^{-1}\,B^* = AA^{\dagger}$$

となっている．また，

$$\begin{aligned} A^{\dagger}A &= C'B'BC = C^*(CC^*)^{-1}(B^*B)^{-1}B^*BC \\ &= C^*(CC^*)^{-1}C \end{aligned}$$

より $(A^{\dagger}A)^* = A^{\dagger}A$ も成り立つ．ゆえに $A^{\dagger}$ は MP 一般逆行列である．

次に一意性を証明する．$A^{\dagger}, Y$ がともに $A$ の MP 一般逆行列であるとする．このとき，

$$\begin{aligned} A^{\dagger} &= A^{\dagger}AA^{\dagger} = A^{\dagger}AYAA^{\dagger} = A^{\dagger}\,(AY)^*\left(AA^{\dagger}\right)^* = A^{\dagger}Y^*A^*A^{\dagger *}A^* \\ &= A^{\dagger}Y^*A^* = A^{\dagger}AY = A^{\dagger}AYAY = \left(A^{\dagger}A\right)^*(YA)^*\,Y \\ &= A^*A^{\dagger *}A^*Y^*Y = A^*Y^*Y = YAY = Y \end{aligned}$$

となっている．■

**例題 15.3** $A=\begin{pmatrix}1 & 1 & 2\\ 1 & 1 & 2\end{pmatrix}$ の MP 一般逆行列を求めよ．

**答．** $\mathrm{rank}(A)=1$ である．まず定理 15.7 の証明にならって，$A$ の階数分解を求める．定理 15.7 の証明中の記号を用いれば

$$A_1=\begin{pmatrix}1\\ 1\end{pmatrix},$$
$$A_2=\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix}=\begin{pmatrix}1\\ 1\end{pmatrix}\begin{pmatrix}1 & 2\end{pmatrix},$$

$Q=(1\quad 2)$ である．したがって，

$$B=\begin{pmatrix}1\\ 1\end{pmatrix},\quad C=\begin{pmatrix}1 & 1 & 2\end{pmatrix}$$

とすると $A=BC$ は階数分解になっている．いま

$$CC^*=\begin{pmatrix}1 & 1 & 2\end{pmatrix}\begin{pmatrix}1\\ 1\\ 2\end{pmatrix}=(6),$$
$$B^*B=\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}=(2)$$

であるから，

$$A^\dagger=\begin{pmatrix}1\\ 1\\ 2\end{pmatrix}\frac{1}{6}\frac{1}{2}\begin{pmatrix}1 & 1\end{pmatrix}=\frac{1}{12}\begin{pmatrix}1 & 1\\ 1 & 1\\ 2 & 2\end{pmatrix}.$$

■

**例題 15.4** $A=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}$ の MP 一般逆行列を求めよ．

**答．** $\mathrm{rank}(A)=2$ である．$P=\begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}$ とすると $AP=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}$ である．定理 15.7 の証明中の記号を用いれば，

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

より $Q = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ である．したがって，

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$\begin{aligned} C &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} P^{-1} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

とすると，$A = BC$ が階数分解である．いま

$$B^*B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$CC^* = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

であるから，

$$A^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

■

**問題 15.2（MacDuffee の定理）**　$A = BC$ を定理 15.7 の分解とする．このとき，$B^*AC^*$ は正則で，

$$A^\dagger = C^*(B^*AC^*)^{-1}B^*$$

であることを示せ．(ヒント：定理 15.8 を利用する．)

ムーア-ペンローズ行列の別の構成方法を述べておく．問題 10.6 を用いれば，$A \in \boldsymbol{K}^{M\times N}$ に対して次のような写像を定義することができる．

$$T_A^{\dagger}(\mathbf{y}) = \begin{cases} \left(T_A|_{\mathcal{R}(A^*)}\right)^{-1}(\mathbf{y}), & \mathbf{y} \in \mathcal{R}(A), \\ \mathbf{0}, & \mathbf{y} \in \mathcal{R}(A)^{\perp}. \end{cases}$$

明らかに $T_A^{\dagger}$ は $\boldsymbol{K}^M$ から $\boldsymbol{K}^N$ への線形写像である．また問題 10.6 より $\mathcal{R}(T_A^{\dagger}) = \mathcal{R}(A^*)$ となっている．$T_A^{\dagger}$ を表現する行列について次のことが成り立つ．

**定理 15.9** $A \in \boldsymbol{K}^{M\times N}$ に対して，$T_A^{\dagger}$ を表現する行列が $A$ の MP 一般逆行列である．すなわち，$T_A^{\dagger} = T_{A^{\dagger}}$ である．

**証明.** $T_A^{\dagger}$ を表現する行列を $B$ とする．$B \in \boldsymbol{K}^{N\times M}$ である．以下 $B = A^{\dagger}$ を証明する．そのため $B$ が定義 15.2 の条件 (1) ～ (4) をみたすことを示す．$\mathbf{y} \in \mathcal{R}(A)$ のとき，ある $\mathbf{z} \in \mathcal{R}(A^*)$ が存在し，$T_A(\mathbf{z}) = T_A|_{\mathcal{R}(A^*)}(\mathbf{z}) = \mathbf{y}$ をみたす．ゆえに

$$T_A T_A^{\dagger}(\mathbf{y}) = T_A\left(T_A|_{\mathcal{R}(A^*)}\right)^{-1}(\mathbf{y}) = T_A(\mathbf{z}) = \mathbf{y}$$

が成り立つ．また，$\mathbf{y} \in \mathcal{R}(A)^{\perp}$ の場合は $T_A T_A^{\dagger}(\mathbf{y}) = T_A(\mathbf{0}) = \mathbf{0}$ である．これより $T_A T_A^{\dagger}$ は $\boldsymbol{K}^M$ から $\mathcal{R}(A)$ への直交射影 $P_{\mathcal{R}(A)}$ となっていることがわかる．$T_A T_A^{\dagger}$ を表現する行列は $AB$ であるから $(AB)^* = AB$ である．また $\mathbf{x} \in \mathcal{N}(A) = \mathcal{R}(A^*)^{\perp}$ のとき，$T_A^{\dagger} T_A \mathbf{x} = \mathbf{0}$ である．そして $\mathbf{x} \in \mathcal{R}(A^*) = \mathcal{R}(T_A^{\dagger})$ のときは $T_A^{\dagger} T_A \mathbf{x} = \left(T_A|_{\mathcal{R}(A^*)}\right)^{-1} T_A \mathbf{x} = \mathbf{x}$ である．ゆえに $T_A^{\dagger} T_A$ は $\boldsymbol{K}^N$ から $\mathcal{R}(T_A^{\dagger})$ への直交射影 $P_{\mathcal{R}(A^*)}$ となっている．したがって，$(BA)^* = BA$ も成り立っている．いま

$$T_A T_A^{\dagger} T_A = P_{\mathcal{R}(A)} T_A = T_A$$

より $ABA = A$ が得られる．また

$$T_A^{\dagger} T_A T_A^{\dagger} = P_{\mathcal{R}(T_A^{\dagger})} T_A^{\dagger} = T_A^{\dagger}$$

であるから $BAB = B$ も成り立っている．よって，定理 15.8 の一意性より $B = A^{\dagger}$ となることがわかる． ■

この定理の証明から次のことがわかる．

**定理 15.10** $A \in \boldsymbol{K}^{M\times N}$ に対して，

$$AA^{\dagger} = P_{\mathcal{R}(A)}, \quad A^{\dagger}A = P_{\mathcal{R}(A^{\dagger})} \tag{15.8}$$

をみたす．また，もし $B \in \boldsymbol{K}^{N\times M}$ が

$$AB = P_{\mathcal{R}(A)}, \quad BA = P_{\mathcal{R}(B)} \tag{15.9}$$

をみたすならば，$B = A^\dagger$ である．

**証明．** 前半の主張は本節のはじめに証明した．後半の主張は，定理 15.9 の証明より $B$ が定理 15.8 の条件をみたしていることよりわかる．■

定義 15.2 は R. ペンローズによる定義である．これとは独立に，E. H. ムーアは定理 15.10 の条件 (15.8) により一般逆行列を定義した．

ムーア-ペンローズ一般逆行列は次章で述べる特異値分解を使って構成することも可能である (定理 16.10 参照)．

**問題 15.3**　$A \in \boldsymbol{K}^{M \times N}$ 行列とする．次のことを示せ．

(1)　$(A^\dagger)^\dagger = A$,

(2)　$(A^\dagger)^* = (A^*)^\dagger$,

(3)　$\lambda \in \boldsymbol{K}, \lambda \neq 0$ のとき $(\lambda A)^\dagger = \lambda^{-1} A^\dagger$.

## 15.3　連立方程式の最小 2 乗解への応用

$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N) \in \boldsymbol{K}^{M \times N}$ とし，$\mathbf{b} \in \boldsymbol{K}^M$ とする．方程式

$$A\mathbf{x} = \mathbf{b} \tag{15.10}$$

を考える．もしこの方程式が解をもたないときには，どのような $\mathbf{x} \in \boldsymbol{K}^N$ に対しても $\|A\mathbf{x} - \mathbf{b}\| > 0$ である．そこで次のような概念を導入する．

**定義 15.3**　もし $\widehat{\mathbf{x}} \in \boldsymbol{K}^N$ が

$$\|A\widehat{\mathbf{x}} - \mathbf{b}\| = \min_{\mathbf{x} \in \boldsymbol{K}^M} \|A\mathbf{x} - \mathbf{b}\| \tag{15.11}$$

をみたすならば，$\widehat{\mathbf{x}}$ を方程式 (15.10) の**最小 2 乗解**という．ここで (15.11) の右辺は $\|A\mathbf{x} - \mathbf{b}\|, \mathbf{x} \in \boldsymbol{K}^M$ の中で最も小さい値を表す記号である．

もし方程式 (15.10) が解 $\mathbf{x}_{sol}$ を持つならば，

$$0 \leq \min_{\mathbf{x} \in \boldsymbol{K}^M} \|A\mathbf{x} - \mathbf{b}\| \leq \|A\mathbf{x}_{sol} - \mathbf{b}\| = 0$$

より，$\mathbf{x}_{sol}$ は (15.10) の最小 2 乗解になっている．最小 2 乗解は解をもたない方程式の場合に役立つ概念である．

ムーア-ペンローズ一般逆行列を用いれば，次のように最小 2 乗解の一つを求め

ることができる．

**定理 15. 11**　$A^\dagger$ を $A$ のムーア-ペンローズ一般逆行列とする．任意の $\mathbf{b} \in \boldsymbol{K}^M$ に対して，$\mathbf{x} = A^\dagger\mathbf{b}$ は

$$A\mathbf{x} = \mathbf{b}$$

の最小 2 乗解になっている．

この定理をいくつかのステップに分けて証明する．はじめに正規方程式を定義する．

**定義 15. 4**　方程式

$$A^* A\mathbf{x} = A^*\mathbf{b} \tag{15.12}$$

を**正規方程式**という．

**補題 15. 12**　$\mathbf{x}' = A^\dagger\mathbf{b}$ とおくと $\mathbf{x}'$ は正規方程式 (15.12) の解になっている．

**証明.** ムーア-ペンローズ一般逆行列の条件 (1) および (3) より次が成り立つ．

$$\begin{aligned} A^* A\mathbf{x}' &= A^* A A^\dagger \mathbf{b} = A^* \left(AA^\dagger\right)^* \mathbf{b} = A^* A^{\dagger *} A^* \mathbf{b} \\ &= \left(AA^\dagger A\right)^* \mathbf{b} = A^*\mathbf{b}. \end{aligned}$$

■

次に正規方程式の解ならば，それは最小 2 乗解であることを示す．

**補題 15. 13**　正規方程式 (15.12) の解は $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解である．

**証明.** $\mathbf{x}$ を (15.12) の解とする．仮定より

$$\begin{pmatrix} \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_N^* \end{pmatrix} A\mathbf{x} = A^* A\mathbf{x} = A^*\mathbf{b} = \begin{pmatrix} \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_N^* \end{pmatrix} \mathbf{b}.$$

したがって，$\mathbf{a}_j^* A\mathbf{x} = \mathbf{a}_j^*\mathbf{b},\ j = 1, \cdots, N$．すなわち

$$0 = \mathbf{a}_j^*(A\mathbf{x} - \mathbf{b}) = \langle A\mathbf{x} - \mathbf{b}, \mathbf{a}_j \rangle$$

が成り立つ．ゆえに，$A\mathbf{x} - \mathbf{b} \in \mathcal{R}(A)^\perp$ である．これより任意の $\mathbf{x}' \in \boldsymbol{K}^M$ に対して，

$$\|A\mathbf{x}' - \mathbf{b}\|^2 = \|A\left(\mathbf{x}' - \mathbf{x}\right) + A\mathbf{x} - \mathbf{b}\|^2$$

$$= \|A(\mathbf{x}' - \mathbf{x})\|^2 + \|A\mathbf{x} - \mathbf{b}\|^2 \geq \|A\mathbf{x} - \mathbf{b}\|^2$$

となるので，$\mathbf{x}$ は $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解になっている．■

**定理 15.11 の証明．** 補題 15.12，補題 15.13 より明らかである．■

以上の証明から次のことがわかる．

**系 15. 14**　正規方程式 (15.12) の解であることと $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解であることは同値である．

**証明．** 正規方程式の解が最小 2 乗解であることは示したので，その逆を証明する．$\mathbf{z}$ を $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解とする．

$$\|A\mathbf{z} - \mathbf{b}\| = \min_{\mathbf{x} \in \boldsymbol{K}^M} \|A\mathbf{x} - \mathbf{b}\| = \min_{\mathbf{y} \in \mathcal{R}(A)} \|\mathbf{y} - \mathbf{b}\|$$

とも表せるので，$A\mathbf{z}$ は $\mathbf{b}$ の $\mathcal{R}(A)$ に関する最良近似になっている．したがって定理 10.21 より $A\mathbf{z} = P_{\mathcal{R}(A)}(\mathbf{b})$ である．$\mathbf{x}' = A^{\dagger}\mathbf{b}$ とおくと，定理 15.11 よりこれも最小 2 乗解であるから，$A\mathbf{x}' = P_{\mathcal{R}(A)}(\mathbf{b})$ である．$A\mathbf{z} = A\mathbf{x}'$ かつ $\mathbf{x}'$ は正規方程式の解であるから (補題 15.12)，

$$A^*A\mathbf{z} = A^*A\mathbf{x}' = A^*\mathbf{b}$$

であり，$\mathbf{z}$ も正規方程式の解である．■

**例 15. 1**　$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ のとき，例題 15.3 より

$$A^{\dagger}\mathbf{b} = \frac{1}{12}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

は $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解である．

最小 2 乗解は一般には一意的には定まらない．ムーア-ペンローズ一般逆行列により作られる最小 2 乗解は，最小 2 乗解の中でノルムが最小になっているものである．すなわち

$$\left\|A^{\dagger}\mathbf{b}\right\| = \min\left\{\|\mathbf{x}'\| : \mathbf{x}' \in \boldsymbol{K}^N \text{ は } A\mathbf{x} = \mathbf{b} \text{ の最小 2 乗解}\right\}$$

が成り立っている (このことは後述の定理 16.11 において証明する．定理 16.11 で

はすべての最小 2 乗解が求められている).

特に行列 $A$ の階数が列の数と一致している場合には，次のことがいえる．

**命題 15.15** $A$ を $(M, N)$ 行列とする．$\mathrm{rank}(A) = N$ ならば $A^*A$ は正則行列であり，$A$ のムーア-ペンローズ一般逆行列は

$$A^\dagger = (A^*A)^{-1} A^*$$

により与えられる．さらに $\mathbf{x}' = A^\dagger \mathbf{b}$ は $A\mathbf{x} = \mathbf{b}$ の一意的な最小 2 乗解となっている．

**証明．** $A^*A$ が正則であることは系 11.13 による．$(A^*A)^{-1} A^*$ がムーア-ペンローズ一般逆行列の条件をすべてみたしていることが容易にわかる．したがって，$\mathbf{x}'$ は $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解である．いま $\mathbf{x}''$ が $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解であるとすると，正規方程式 (15.12) の解でもあるから，$A^*A\mathbf{x}'' = A^*\mathbf{b}$ であり，したがって

$$\mathbf{x}'' = (A^*A)^{-1} A^*A\mathbf{x}'' = (A^*A)^{-1} A^*\mathbf{b} = A^\dagger \mathbf{b} = \mathbf{x}'$$

となっている．■

**例題 15.5** $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ のとき，$A\mathbf{x} = \mathbf{b}$ の最小 2 乗解を求めよ．

**答．** $A^* = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ であるから，$A^*A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$，また $A^*\mathbf{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ である．最小 2 乗解は

$$\mathbf{x} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

■

## 15.4 データの直線，曲線による当てはめへの応用

次のような数学的な問題を考えてみよう．いま平面上に $N$ 個の点 $(x_1, y_1), \cdots, (x_N, y_N)$ が与えられたとする．直線 $l : y = \alpha_1 + \alpha_2 x$ を適当にひき，

$$\varepsilon_j(\alpha_1, \alpha_2) = y_j - (\alpha_1 + \alpha_2 x_j)$$

とおく．$\varepsilon_j(\alpha_1, \alpha_2)$ は点 $(x_j, y_j)$ と直線 $y = \alpha_1 + \alpha_2 x$ 上の点 $(x_j, \alpha_1 + \alpha_2 x_j)$ と

の差を表している (図 15.1 参照)．この差の平方和

$$\varepsilon(\alpha_1, \alpha_2) = \sum_{j=1}^{N} |\varepsilon_j(\alpha_1, \alpha_2)|^2$$

をとる．$\varepsilon(\alpha_1, \alpha_2)$ は $\alpha_1, \alpha_2$ の関数になっている．そこで $\varepsilon(\alpha_1, \alpha_2)$ が最も小さくなるような $a_1, a_2$, すなわち

$$\varepsilon(a_1, a_2) = \min_{\alpha_1, \alpha_2 \in \boldsymbol{R}} \varepsilon(\alpha_1, \alpha_2)$$

をみたす $a_1, a_2$ を求めることを考える．このような $a_1, a_2$ を係数とする直線 $y = a_1 + a_2 x$ をデータ $(x_1, y_1), \cdots, (x_N, y_N)$ にフィットした**最小 2 乗直線**あるいは**重回帰直線**という．この直線は，たとえばデータ $(x_1, y_1), \cdots, (x_N, y_N)$ が与えられたときに $x_{N+1}$ に対する未知の値 $y_{N+1}$ を知りたいときに用いられることがある．$\widehat{y}_{N+1} = a_1 + a_2 x_{N+1}$ を予測値とするのである．このような方法は**重回帰分析**と呼ばれている．

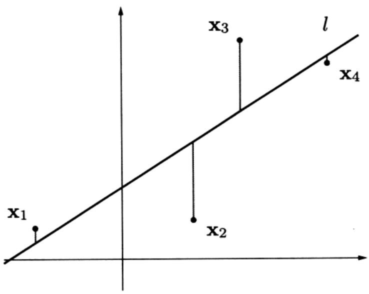



図 15.1

ここでは問題を線形代数の用語を用いて定式化して調べていく．$\mathbf{y} = (y_j)$, $\mathbf{x} = (x_j) \in \boldsymbol{R}^N$ とする．これらは与えられたデータである．

$$X = \begin{pmatrix} \mathbf{1}_N & \mathbf{x} \end{pmatrix} \in \boldsymbol{R}^{N \times 2}$$

とおく．$\mathbf{a} = (a_j)_{j=1,2} \in \boldsymbol{R}^2$ に対して

$$X\mathbf{a} - \mathbf{y} = \begin{pmatrix} a_1 + a_2 x_1 - y_1 \\ \vdots \\ a_1 + a_2 x_N - y_N \end{pmatrix}$$

であるから，$\varepsilon(a_1, a_2)$ を最小にする $\mathbf{a} = (a_j)_{j=1,2} \in \boldsymbol{R}^2$ は，$X\mathbf{a} = \mathbf{y}$ の最小 2 乗解にほかならない．したがって，たとえば

$$\mathbf{a} = X^{\dagger}\mathbf{y}$$

を係数とする直線 $y = a_1 + a_2 x$ がデータにフィットした直線である．応用上は $\mathrm{rank}(X) = 2$ であることが多いが，この場合は

$$\mathbf{a} = X^{\dagger}\mathbf{y} = (X^* X)^{-1} X^* \mathbf{y}$$

により求める係数が得られる．

**例題 15.6** $(1,1)$, $\left(3, \dfrac{3}{2}\right)$, $(4,5)$, $\left(6, \dfrac{11}{2}\right)$ にフィットする最小 2 乗直線を求めよ．

**答.**

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ \dfrac{3}{2} \\ 5 \\ \dfrac{11}{2} \end{pmatrix}$$

とすると

$$X^T X = \begin{pmatrix} 4 & 14 \\ 14 & 62 \end{pmatrix}, \quad X^T \mathbf{y} = \begin{pmatrix} 13 \\ \dfrac{117}{2} \end{pmatrix}.$$

ここで

$$\begin{pmatrix} 4 & 14 \\ 14 & 62 \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{31}{26} & -\dfrac{7}{26} \\ -\dfrac{7}{26} & \dfrac{1}{13} \end{pmatrix}$$

であるから，正規方程式 $X^T X\mathbf{a} = X^T\mathbf{y}$ の解は

$$\mathbf{a} = \begin{pmatrix} \dfrac{31}{26} & -\dfrac{7}{26} \\ -\dfrac{7}{26} & \dfrac{1}{13} \end{pmatrix} \begin{pmatrix} 13 \\ \dfrac{117}{2} \end{pmatrix} = \begin{pmatrix} -\dfrac{1}{4} \\ 1 \end{pmatrix}$$

である．ゆえに求める直線は

$$y = -\frac{1}{4} + x$$

である．■

直線がどれだけ与えられたデータにフィットしているかを示す量として，次の

**重相関係数**がある．これは $\widehat{\mathbf{y}} = X\mathbf{a}$, $\overline{y} = \dfrac{1}{N}\sum_{j=1}^{N} y_j$ として

$$R = \left( \frac{\|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2}{\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2} \right)^{\frac{1}{2}}$$

により定義される．第 17.2 節で証明することであるが, $0 \leq R \leq 1$ であり，$R = 1$ ならば $\mathbf{y} = \widehat{\mathbf{y}}$ である．例題 15.6 の場合，$R = \sqrt{\dfrac{4}{5}} \fallingdotseq 0.9$ である．

**問題 15. 4**　$R = \sqrt{\dfrac{4}{5}}$ を示せ．

さて，たとえば $(1, 8)$, $(2, 3)$, $(3, 0)$, $(4, -1)$, $(5, 0)$ というデータが得られたとき，これにフィットする最小 2 乗直線は,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (X^T X)^{-1} X^T \mathbf{y}$$

によって与えられ, $y = 8 - 2x$ である．この重相関係数を計算すると $R = \sqrt{\dfrac{40}{54}} \fallingdotseq 0.86$ である．じつはここで考えている点は 2 次関数 $y = (x-4)^2 - 1$ のグラフ上の点になっている．

**問題 15. 5**　$R = \sqrt{\dfrac{40}{54}}$ を示せ．

これまではデータにフィットする直線を求める方法を述べたが, 同様の方法でデータにフィットする $p$ 次曲線を求めることもできる．データ $(x_1, y_1), \cdots, (x_N, y_N)$ と $a_1, \cdots, a_{p+1}$ に対して,

$$\varepsilon_j\,(a_1, \cdots, a_{p+1}) = y_j - \left(a_1 + a_2 x + a_3 x_j^2 + \cdots + a_{p+1} x_j^p\right)$$

とおき,

$$\varepsilon(a_1, \cdots, a_{p+1}) = \sum_{j=1}^{N} \left|\varepsilon_j\,(a_1, \cdots, a_{p+1})\right|^2$$

とする．このとき問題は,

$$\varepsilon(a_1, \cdots, a_{p+1}) = \min_{b_1, \cdots, b_{p+1} \in \boldsymbol{R}} \varepsilon(b_1, \cdots, b_{p+1})$$

をみたす $a_1, \cdots, a_{p+1} \in \boldsymbol{R}$ を求めることになる．これを線形代数の用語を用いて記述すれば，次のようになる．

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^p \end{pmatrix}$$

とおく．このとき正規方程式

$$X^T X\mathbf{a} = X^T\mathbf{y}$$

をみたす $\mathbf{a} = (a_j) \in \boldsymbol{R}^{p+1}$ に対して，$f(x) = a_1 + a_2x + \cdots + a_{p+1}x^p$ が求める $p$ 次関数である．これをデータ $(x_1, y_1), \cdots, (x_N, y_N)$ にフィットした**最小 2 乗 $p$ 次曲線**という．

なお補間多項式の項で述べたが，$x_j \neq x_k\ (j \neq k)$ をみたしているとき，ちょうど $f(x_j) = y_j\ (j = 1, \cdots, N)$ をみたす $N-1$ 次関数 $f(x)$ が存在する．実際，$X^T$ はファンデルモンド行列になるので，$\left(X^T\right)^{-1}$, $X^{-1}$ が存在し，正規方程式の解は $X\mathbf{a} = \mathbf{y}$ をみたす．すなわち $R = 1$ となる．またこの場合の最小 2 乗 $N-1$ 次曲線は補間多項式により与えられる．

ここでは $p < N-1$ の場合に興味がある．

**問題 15.6**　$(1, 8), (2, 3), (3, 1), (4, -1), (5, 0)$ の最小 2 乗 2 次曲線を求めよ．

本節での考察は第 17.2 節で，もう少し一般的な立場から論ずる．

## 15.5　種々の一般逆行列

これまで一般逆行列，ムーア-ペンローズ一般逆行列について学んできた．一般逆行列に比べて，ムーア-ペンローズ一般逆行列は多くの付加条件を課している．しかし，用途によってはムーア-ペンローズ一般逆行列よりも少ない条件をみたす一般逆行列を考えると便利なこともある．まず定義 15.2 で述べた条件を振り返っておこう．$A \in \boldsymbol{K}^{M\times N}$ とする．$X \in \boldsymbol{K}^{N\times M}$ に対して，

(1)　$AXA = A$,

(2)　$XAX = X$,

(3)　$(AX)^* = AX$,

(4)　$(XA)^* = XA$

をそれぞれペンローズ条件 $(j)$, $j = 1, 2, 3, 4$ という．

**定義 15.5**　ペンローズ条件 $(i), \cdots, (k)$ をみたすものを $A$ の $\{i, \cdots, k\}$**-一般**

**逆行列**という．$A$ の $\{i,\cdots k\}$-一般逆行列全体のなす集合を $A\{i,\cdots,k\}$ と表す．

$A$ の $\{1\}$-一般逆行列が一般逆行列である．また $A$ の $\{1,2,3,4\}$-一般逆行列がムーア-ペンローズ一般逆行列 $A^\dagger$ である．

**定義 15.6**　$A$ の $\{1,2\}$-一般逆行列を $A$ の**反射型一般逆行列**という．
$A$ の $\{1,3\}$-一般逆行列を $A$ の**最小 2 乗型一般逆行列**という．
$A$ の $\{1,4\}$-一般逆行列を $A$ の**ノルム最小型一般逆行列**という．

定義からムーア-ペンローズ一般逆行列は反射型，最小 2 乗型，かつノルム最小型一般逆行列になっている．

定義 15.6 で定義した一般逆行列について，基本的な性質と名前の由来を述べておく．

### 15.5.1　反射型一般逆行列

$A$ の一般逆行列 $A^-$ について，一般には $A^-$ が $A$ の一般逆行列になっているとは限らない．これに対して，$A^-$ が $A$ の一般逆行列であり，かつ $A$ が $A^-$ の一般逆行列になっているものが，反射型一般逆行列である．

**問題 15.7**　$B,C$ が $A$ の一般逆行列ならば，$BAC$ は $A$ の反射型一般逆行列であることを示せ．

反射型一般逆行列は次のように特徴付けられる．

**補題 15.16**　$A^-\in A\{1\}$ とする．次の (i), (ii) は同値である．

(i)　$A^-\in A\{1,2\}$,

(ii)　$\mathrm{rank}(A)=\mathrm{rank}(A^-)$.

**証明.** (ii) $\Rightarrow$ (i)　$A^-=(I-A^-A)\,A^-+A^-AA^-$ であり，

$$R\left(I-A^-A\right)\cap R\left(A^-A\right)=\{0\}$$

である．実際，$\mathbf{y}$ が左辺に属していれば，$\mathbf{y}=(I-A^-A)\,\mathbf{x}=A^-A\mathbf{z}$ をみたす $\mathbf{x},\mathbf{z}$ が存在する．ゆえに

$$A\mathbf{y}=A\mathbf{x}-AA^-A\mathbf{x}=A\mathbf{x}-A\mathbf{x}=\mathbf{0},$$
$$A\mathbf{y}=AA^-A\mathbf{z}=A\mathbf{z}.$$

したがって $A\mathbf{z}=\mathbf{0}$．これより $\mathbf{y}=A^-A\mathbf{z}=\mathbf{0}$ が導かれる．ゆえに

$$\operatorname{rank}(A^-) = \operatorname{rank}(\left(I - A^- A\right) A^-) + \operatorname{rank}\left(A^- A A^-\right). \tag{15.13}$$

いま

$$\begin{aligned}\operatorname{rank}(A) &\geq \operatorname{rank}(A^- A A^-) \geq \operatorname{rank}(A^- A A^- A) \\ &= \operatorname{rank}\left(A^- A\right) = \operatorname{rank}(A)\end{aligned} \tag{15.14}$$

より $\operatorname{rank}(A^- A A^-) = \operatorname{rank}(A) = \operatorname{rank}(A^-)$ である．したがって (15.13) より $\operatorname{rank}((I - A^- A)\, A^-) = 0$，すなわち $A^- = A^- A A^-$ が得られる．

(i) ⇒ (ii)　(15.14) は $A^- \in A\{1\}$ という仮定のみから導かれるものであることに注意する．$A^- \in A\{2\}$ により $(I - A^- A)A^- = O$ であるから，

$$\operatorname{rank}(A) = \operatorname{rank}(A^- A A^-) = \operatorname{rank}\left(A^-\right)$$

である．■

### 15.5.2　最小 2 乗型一般逆行列

定理 15.11 の証明を想起すると，方程式 $A\mathbf{x} = \mathbf{b}$ に対して，$A^\dagger \mathbf{b}$ が最小 2 乗解であることを証明するのに，ムーア-ペンローズ一般逆行列の条件のうち (1) と (3) のみを用いていることがわかる．すなわちムーア-ペンローズ一般逆行列でなくとも，一般の $\{1,3\}$-一般逆行列 $A^{(1,3)}$ に対して $A^{(1,3)}\mathbf{b}$ が $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解になっていることがわかる．これが $\{1,3\}$-一般逆行列を最小 2 乗型一般逆行列と呼ぶ理由である．

最小 2 乗型一般逆行列の例としては次のようなものがある．

**定理 15.17**　$(A^* A)^- A^*$ は $A$ の最小 2 乗型一般逆行列である．したがって $(A^* A)^- A^* \mathbf{b}$ は (15.10) の最小 2 乗解になっている．

**証明.** $X = (A^* A)^- A^*$ とおく．定理 11.12 の証明より，$\mathcal{R}(A^* A) = \mathcal{R}(A^*)$ であるから，ある行列 $C$ で，$A^* = A^* A C$ をみたすものが存在する (たとえば問題 10.7 参照)．したがって，$A = C^* A^* A$ であるから，

$$AXA = C^* A^* A\,(A^* A)^-\, A^* A = C^* A^* A = A$$

である．ゆえに $X \in A\{1\}$ となっている．次に

$$\begin{aligned}AX &= C^* A^* A\,(A^* A)^-\, A^* = C^* A^* A\,(A^* A)^-\, A^* A C \\ &= C^* A^* A C = (AC)^*\,(AC).\end{aligned}$$

これより $(AX)^* = AX$ を得る．■

### 15.5.3 ノルム最小型一般逆行列

$A \in \boldsymbol{K}^{M \times N}$, $\mathbf{b} \in \boldsymbol{K}^{M}$ とする．もし方程式 $A\mathbf{x} = \mathbf{b}$ が解を持ったとすると，定理 15.3 より，$A$ の任意の一般逆行列 $A^{-}$ に対して，$A^{-}\mathbf{b}$ は $A\mathbf{x} = \mathbf{b}$ の解になっている．ノルム最小型一般逆行列に対して次のことが成り立っている．

**定理 15.18**　方程式 $A\mathbf{x} = \mathbf{b}$ が解を持つとする．$A^{(1,4)}$ を $A$ のノルム最小型一般逆行列とする．このとき $A$ の任意の一般逆行列 $A^{-}$ に対して

$$\left\|A^{(1,4)}\mathbf{b}\right\| \leq \left\|A^{-}\mathbf{b}\right\|$$

が成り立つ．

**証明.** $\mathbf{y}$ を $A\mathbf{x} = \mathbf{b}$ の解とする．$P = A^{(1,4)}A$ とおくと，

$$PP = A^{(1,4)}AA^{(1,4)}A = A^{(1,4)}A = P$$

であり，また $P^{*} = P$ である．定理 10.20 より $P$ は $\mathcal{R}(P)$ への直交射影行列である．ゆえに定理 10.20 より

$$\left\|PA^{-}A\mathbf{y}\right\| \leq \left\|A^{-}A\mathbf{y}\right\| = \left\|A^{-}\mathbf{b}\right\|.$$

一方，

$$PA^{-}A = A^{(1,4)}AA^{-}A = A^{(1,4)}A = P$$

であるから，

$$\left\|PA^{-}A\mathbf{y}\right\| = \|P\mathbf{y}\| = \left\|A^{(1,4)}A\mathbf{y}\right\| = \left\|A^{(1,4)}\mathbf{b}\right\|$$

である．■

ムーア-ペンローズ一般逆行列はノルム最小型一般逆行列であるが，次のようにノルム最小型一般逆行列を作ることもできる．

**定理 15.19**　$A^{*}\left(AA^{*}\right)^{-}$ は $A$ のノルム最小型一般逆行列である．

**証明.** $X = A^{*}\left(AA^{*}\right)^{-}$ とおく．$\mathcal{R}(AA^{*}) = \mathcal{R}(A)$ より $AA^{*}C = A$ をみたす行列 $C$ が存在する (定理 11.12，問題 10.7 参照)．したがって，

$$AXA = AA^{*}\left(AA^{*}\right)^{-}AA^{*}C = AA^{*}C = A$$

であるから $X \in A\{1\}$ である．また $A^{*} = C^{*}AA^{*}$ であるから，

$$\begin{aligned} XA &= A^{*}\left(AA^{*}\right)^{-}AA^{*}C = C^{*}AA^{*}\left(AA^{*}\right)^{-}AA^{*}C \\ &= C^{*}AA^{*}C = \left(A^{*}C\right)^{*}\left(A^{*}C\right) \end{aligned}$$

より $(XA)^* = XA$ である．よって $X$ がノルム最小型一般逆行列であることが得られる． ■

# 第16章　特異値分解とその応用

本章では行列の特異値分解，特異値標準化とその応用について述べる．

**ポイント16**　特異値分解あるいは特異値標準化は非常に応用に富んでいる．たとえば各種一般逆行列は特異値標準形の形で特徴付けられる．また特異値分解の主成分分析への応用もある．本章では特異値分解が具体的に計算できるようにしてほしい．なお特異値を計算機で計算するソフトウェアなどもあるので，必要に応じて利用すると良いだろう．

## 16.1　行列の特異値分解

この節では行列の特異値標準化，特異値分解について述べる．そのためにいくつかの補題を準備しておく．次の補題は補題 13.8 (1) として述べたものである．

**補題 16.1**　$A$ を $(M,N)$ 行列とする．このとき，$A^*A$, $AA^*$ はそれぞれ $N$ 次および $M$ 次のエルミート行列であり，その固有値はすべて非負である．

次の補題は特異値の定義に用いられる．

**補題 16.2**　$A$ を $(M,N)$ 行列とし，$r = \mathrm{rank}(A)$ とする．このとき，$A^*A$ の 0 でない固有値は重複度もいれて $r$ 個存在し，それらは $AA^*$ の 0 でない固有値と重複度もいれて一致する．

**証明.**　定理 11.12 より，$\mathrm{rank}(AA^*) = \mathrm{rank}(A^*A) = \mathrm{rank}(A)$ より，$AA^*$ と $A^*A$ の 0 でない固有値の個数は，重複度をこめて $r$ 個存在する．いま $\lambda \neq 0$ を $A^*A$ の固有値とし，$\mathbf{u} \in \boldsymbol{K}^N$ をその固有ベクトルとする．このとき，$AA^*A\mathbf{u} = \lambda A\mathbf{u}$ である．ここで $A\mathbf{u} \neq \mathbf{0}$ である．なぜなら，もし $A\mathbf{u} = \mathbf{0}$ ならば $A^*A\mathbf{u} = \mathbf{0}$ と

なり，$\mathbf{u}$ が $\lambda \neq 0$ の固有ベクトルであることに反するからである．したがって，$\lambda$ は $A\mathbf{u}$ を固有ベクトルとする $AA^*$ の固有値となっている．$A$ と $A^*$ の役割を入れ替えれば，$AA^*$ の 0 でない固有値が $A^*A$ の固有値であることも示される．いま $\mathbf{u}_1, \cdots, \mathbf{u}_s$ を $A^*A$ の固有値 $\lambda \neq 0$ の固有空間 $V$ の基底とする．このとき $A\mathbf{u}_1, \cdots, A\mathbf{u}_s$ は一次独立である．なぜならば，$\sum_{j=1}^{s} \alpha_j A\mathbf{u}_j = \mathbf{0}$ とすると

$$\lambda \sum_{j=1}^{s} \alpha_j \mathbf{u}_j = \sum_{j=1}^{s} \alpha_j A^* A\mathbf{u}_j = \mathbf{0}$$

となるからである．したがって $A\mathbf{u}_1, \cdots, A\mathbf{u}_s$ は $AA^*$ の $\lambda$ に関する固有ベクトルで 1 次独立になっている．ゆえに $AA^*$ の $\lambda$ に関する固有空間を $W$ とすると $\dim V \leq \dim W$ である．$A$ と $A^*$ の役割を代えて議論すれば $\dim W \leq \dim V$ も得られる． ■

**定義 16.1** $A$ を $(M, N)$ 行列で，$r = \mathrm{rank}(A)$, $L = \min\{M, N\}$ とする．$A^*A$ の重複度までいれた 0 でない固有値を $\lambda_1, \cdots, \lambda_r$ とする．

$$\sigma_j = \sqrt{\lambda_j} \quad (j = 1, \cdots, r)$$

とする．$\sigma_1, \cdots, \sigma_r$ を $A$ の**特異値**という．本書では便宜上，$r < L$ の場合には

$$\sigma_k = 0 \quad (k = r+1, \cdots, L)$$

とおき，$\sigma_1, \cdots, \sigma_r, \sigma_{r+1}, \cdots, \sigma_L$ を $A$ の特異値という．また特異値は

$$\sigma_1 \geq \cdots \geq \sigma_L$$

のように番号をつけるものとする．$1 \leq j \leq r$ のとき，$\lambda_j$ の重複度によって $\sigma_j$ の重複度を定義し，$r < L$ の場合は $L - r$ を特異値 0 の重複度とする．

次の定理は基本的なものである．

**定理 16.3** $A$ を $(M, N)$ 行列とし，$M \geq N$ とする．$\sigma_1, \cdots, \sigma_N$ を $A$ の特異値とする．このとき，$M$ 次ユニタリー行列 $U$ と $N$ 次ユニタリー行列 $V$ が存在し，

$$U^* A V = \begin{pmatrix} \sigma_1 & & O \\ & \ddots & \\ O & & \sigma_N \\ & O & \end{pmatrix} \tag{16.1}$$

とできる.

**証明.** $A^*A$ は非負定値エルミート行列であるから，ある $N$ 次ユニタリー行列 $V$ が存在し，

$$V^*(A^*A)V = \begin{pmatrix} \sigma_1^2 & & O \\ & \ddots & \\ O & & \sigma_N^2 \end{pmatrix}$$

とできる．$V = (\mathbf{v}_1, \cdots, \mathbf{v}_N)$ とすると，$\{\mathbf{v}_1, \cdots \mathbf{v}_N\}$ は $\boldsymbol{K}^N$ の正規直交基底である．$r = \mathrm{rank}(A^*A)$ とする．$r < N$ の場合，

$$\sigma_1^2 \geq \cdots \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_N^2 = 0$$

となっている．$r = N$ の場合は

$$\sigma_1^2 \geq \cdots \geq \sigma_r^2 = \sigma_N^2 > 0$$

が成り立っている．

$$\mathbf{u}_j = \frac{1}{\sigma_j} A\mathbf{v}_j \quad (j = 1, \cdots, r)$$

とおく．このとき $1 \leq i, j \leq r$ に対して，

$$\begin{aligned} \langle \mathbf{u}_i, \mathbf{u}_j \rangle &= \frac{1}{\sigma_i \sigma_j} \langle A\mathbf{v}_i, A\mathbf{v}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle \mathbf{v}_i, A^* A\mathbf{v}_j \rangle \\ &= \frac{\sigma_j^2}{\sigma_i \sigma_j} \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \frac{\sigma_j^2}{\sigma_i \sigma_j} \delta_{ij} \\ &= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \end{aligned}$$

が成り立つ．ゆえに $\mathbf{u}_1, \cdots, \mathbf{u}_r$ は $\boldsymbol{K}^M$ の正規直交系である．もし $r < M$ ならば，ある $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_M \in \boldsymbol{K}^M$ を加えて，$\{\mathbf{u}_1, \cdots, \mathbf{u}_M\}$ が $\boldsymbol{K}^M$ の正規直交基底となるようにできる．$U = (\mathbf{u}_1, \cdots, \mathbf{u}_M)$ はユニタリー行列である．さらに

$$\begin{aligned} U^*AV &= U^*A(\,\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_N\,) \\ &= \begin{pmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_r^* \\ \mathbf{u}_{r+1}^* \\ \vdots \\ \mathbf{u}_M^* \end{pmatrix} \begin{pmatrix} A\mathbf{v}_1 & \cdots & A\mathbf{v}_r & A\mathbf{v}_{r+1} & \cdots & A\mathbf{v}_N \end{pmatrix} \end{aligned}$$

$$
= \begin{pmatrix}
\langle A\mathbf{v}_1, \mathbf{u}_1\rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_1\rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_1\rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_1\rangle \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\langle A\mathbf{v}_1, \mathbf{u}_r\rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_r\rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_r\rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_r\rangle \\
\langle A\mathbf{v}_1, \mathbf{u}_{r+1}\rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_{r+1}\rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_{r+1}\rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_{r+1}\rangle \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\langle A\mathbf{v}_1, \mathbf{u}_M\rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_M\rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_M\rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_M\rangle
\end{pmatrix}
$$

となっている．ここで，$1 \leq i, j \leq r$ のときは

$$
\langle A\mathbf{v}_j, \mathbf{u}_i\rangle = \sigma_j \langle \mathbf{u}_j, \mathbf{u}_i\rangle = \sigma_j \delta_{ij}
$$

である．また $r+1 \leq i \leq M$, $1 \leq j \leq r$ の場合は

$$
\langle A\mathbf{v}_j, \mathbf{u}_i\rangle = \sigma_j \langle \mathbf{u}_j, \mathbf{u}_i\rangle = 0.
$$

$\mathbf{v}_j$ が $A^*A$ の $\sigma_j^2$ に関する固有ベクトルであることに注意すれば，$r+1 \leq j \leq N$ のとき，

$$
\|A\mathbf{v}_j\|^2 = \langle A\mathbf{v}_j, A\mathbf{v}_j\rangle = \langle \mathbf{v}_j, A^*A\mathbf{v}_j\rangle = \sigma_j^2 \langle \mathbf{v}_j, \mathbf{v}_j\rangle = 0
$$

を得る．よって (16.1) が成り立つ．■

**定義 16.2**　定理 16.3 において，(16.1) を $A$ の**特異値標準化**といい，(16.1) の右辺を $A$ の**特異値標準形**という．

特異値は $A$ により一意的に定まるが，特異値標準形を与えるユニタリー行列 $U, V$ は一意的ではない (後述の例題 16.1 参照)．

定理 16.3 では $M \geq N$ を仮定しているが，$N \geq M$ の場合も同様の結果が成り立つことを示す．次の補題から始める．

**補題 16.4**　$A$ の特異値と $A^*$ の特異値は重複度をこめて一致する．

**証明．** 補題 16.2 より，$A^*A$ と $(A^*)^*A^* = AA^*$ の 0 でない固有値は重複度もいれて一致するので，$A$ と $A^*$ の特異値は重複度もいれて一致する．■

**定理 16.5**　$A$ を $(M, N)$ 行列とし，$N \geq M$ とする．$\sigma_1, \cdots, \sigma_M$ を $A$ の特異値とする．このとき $M$ 次ユニタリー行列 $U$ と $N$ 次ユニタリー行列 $V$ が存在し，

$$U^*AV = \begin{pmatrix} \sigma_1 & & O & \\ & \ddots & & O \\ O & & \sigma_M & \end{pmatrix} \tag{16.2}$$

とできる．

**証明．** 補題 16.4 より，$A^*$ の特異値は $A$ の特異値と重複度もいれて一致する．$A^*$ は $(N, M)$ 行列で，$N \geq M$ であるから，定理 16.3 より，ある $N$ 次ユニタリー行列 $V$ と $M$ 次ユニタリー行列 $U$ が存在し，

$$V^*A^*U = \begin{pmatrix} \sigma_1 & & O \\ & \ddots & \\ O & & \sigma_M \\ & O & \end{pmatrix}.$$

したがって，両辺の共役をとれば定理が得られる．■

**定義 16.3**　$A \in \boldsymbol{K}^{M \times N}$, $N \geq M$ の場合は (16.2) を $A$ の**特異値標準化**といい，(16.2) の右辺を $A$ の**特異値標準形**という．

**問題 16.1**　定理 16.5 を定理 16.3 の証明方法を用いて直接証明せよ．

定理 16.3 と定理 16.5 をまとめて，記法 (2.9) を用いて記述すると次のようになる．

**定理 16.6**　$A$ を $(M, N)$ 行列とし，$\sigma_1, \cdots, \sigma_L$ $(L = \min\{M, N\})$ を $A$ の特異値とする．このときある $M$ 次ユニタリー行列とある $N$ 次ユニタリー行列 $V$ を

$$U^*AV = \mathrm{diag}_{M,N}(\sigma_1, \cdots, \sigma_L) \tag{16.3}$$

となるようにとれる．

この定理において，(16.3) の右辺 $\mathrm{diag}_{M,N}(\sigma_1, \cdots, \sigma_L)$ を $A$ の**特異値標準形**という．

**系 16.7（特異値分解）**　$A$ を $(M, N)$ 行列とする．$\sigma_1, \cdots, \sigma_L$ を $A$ の特異値とする．このとき，$\boldsymbol{K}^M$ の正規直交基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ と $\boldsymbol{K}^N$ の正規直交基底 $\mathbf{v}_1, \cdots, \mathbf{v}_N$ が存在し，

$$A = \sum_{j=1}^{L} \sigma_j \, \mathbf{u}_j \, \mathbf{v}_j^*$$

が成り立っている．

**証明．** $M \geq N$ の場合，定理 16.3 より

$$A = UU^*AVV^* = U \begin{pmatrix} \sigma_1 & & O \\ & \ddots & \\ O & & \sigma_N \\ & O & \end{pmatrix} V^*$$

$$= (\mathbf{u}_1, \cdots, \mathbf{u}_M) \begin{pmatrix} \sigma_1 & & O \\ & \ddots & \\ O & & \sigma_N \\ & O & \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_N^* \end{pmatrix}$$

$$= \sum_{j=1}^{N} \sigma_j \, \mathbf{u}_j \, \mathbf{v}_j^*$$

が成り立つ．$N \geq M$ の場合も定理 16.5 より同様に示される． ■

**例題 16.1** $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ の特異値標準形と特異値分解を求めよ．

**答．** $A^*A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $\Delta_{A^*A}(z) = (z-2)^2$ であるから，固有値は 2 である．したがって $A$ の特異値は $\sigma_1 = \sigma_2 = \sqrt{2}$ である．$2I - A^*A = O$ であるから，$x_2 = s$, $x_1 = t$ が $(2I - A^*A)\,\mathbf{x} = 0$ の解である．ゆえに 2 の固有空間 $V_2$ の正規直交基底は，

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

である．$V = (\mathbf{v}_1, \mathbf{v}_2)$ とおく．次に

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} A\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} A\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

とし，$U=(\mathbf{u}_1,\mathbf{u}_2)$ とする．求める特異値標準形は

$$U^*AV=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}\sqrt{2} & 0\\ 0 & \sqrt{2}\end{pmatrix}$$

である．また，特異値分解は，

$$\begin{aligned}A&=\sqrt{2}\mathbf{u}_1\mathbf{v}_1^*+\sqrt{2}\mathbf{u}_2\mathbf{v}_2^*\\ &=\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}+\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix}\end{aligned}$$

である．このほかにたとえば，

$$\frac{1}{\sqrt{2}}\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}=\begin{pmatrix}\sqrt{2} & 0\\ 0 & \sqrt{2}\end{pmatrix}$$

なども特異値標準形を与える．■

**問題 16. 2**　次の行列の特異値標準形と特異値分解を求めよ．

(1) $\begin{pmatrix}1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1\\ 0 & 1 & 1 & 0\\ 1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1\end{pmatrix}$　(2) $\begin{pmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{pmatrix}$　(3) $\begin{pmatrix}1 & -1\\ -2 & 2\\ 2 & -2\end{pmatrix}$

本節の最後に，特異値標準化の一つの応用例として，極分解と呼ばれる行列の分解公式を証明する．複素数の場合，複素数 $z$ は $z=|z|\,e^{i\theta}$ のように，$z$ を非負の数 $|z|$ と円周上の点 $e^{i\theta}$ に分解できる．この分解の行列版が行列の極分解である．

**定理 16. 8（極分解）**　$A\in\boldsymbol{K}^{N\times N}$ とする．このとき，ある非負定値行列 $P_1,P_2\in\boldsymbol{K}^{N\times N}$ とユニタリー行列 $U_1,U_2\in\boldsymbol{K}^{N\times N}$ により

$$A=P_1U_1=U_2P_2$$

と表せる．この分解においては $P_1,P_2$ の固有値はそれぞれ $A$ の特異値と重複度もこめて一致している．

**証明．** 定理16.6よりあるユニタリー行列 $Q,V$ で，$Q^*AV=\mathrm{diag}_{N,N}(\sigma_1,\cdots,\sigma_N)$ となるものが存在する．このとき，

$$A=Q\mathrm{diag}_{N,N}(\sigma_1,\cdots,\sigma_N)\,V^*=Q\mathrm{diag}_{N,N}(\sigma_1,\cdots,\sigma_N)\,Q^*QV^*$$

である．$P_1 = Q\mathrm{diag}_{N,N}(\sigma_1, \cdots, \sigma_N)Q^*$, $U_1 = QV^*$ とすれば，$P_1$ は非負定値であり，$U_1$ はユニタリーである．また

$$A = QV^*V\mathrm{diag}_{N,N}(\sigma_1, \cdots, \sigma_N)V^*$$

とし，$P_2 = V\mathrm{diag}_{N,N}(\sigma_1, \cdots, \sigma_N)V^*$, $U_2 = QV^*$ とすれば，定理の後半の分解も証明できる．■

正方行列ではない一般の行列に対しては次の定理が成り立つ．

**定理 16.9** $A \in \boldsymbol{K}^{M\times N}$ とする．

(1) $M \leq N$ の場合，ある $U \in \boldsymbol{K}^{M\times N}$ で $U^*U = I_N$ をみたすものとある非負定値行列 $P \in \boldsymbol{K}^{M\times M}$ が存在し，

$$A = PU$$

とできる．またこの分解において $P$ の固有値は $A$ の特異値と重複度をこめて一致している．

(2) $N \leq M$ の場合，ある $U \in \boldsymbol{K}^{M\times N}$ で $U^*U = I_N$ をみたすものとある非負定値行列 $P \in \boldsymbol{K}^{N\times N}$ が存在し，

$$A = UP$$

とできる．この分解において $P$ の固有値は $A$ の特異値と重複度をこめて一致している．

**証明．** まず (2) を示す．$A \in \boldsymbol{K}^{M\times N}$, $N \leq M$ の場合，定理 16.6 よりあるユニタリー行列 $Q \in K^{M\times M}$, $V \in K^{N\times N}$ で，$Q^*AV = \mathrm{diag}_{M,N}(\sigma_1, \cdots, \sigma_N)$ となるものが存在する．$Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}$, $Q_1 \in K^{M\times N}$, $Q_2 \in K^{M\times(M-N)}$ とおく．このとき

$$I_M = Q^*Q = \begin{pmatrix} Q_1^* \\ Q_2^* \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} = \begin{pmatrix} Q_1^*Q_1 & Q_1^*Q_2 \\ Q_2^*Q_1 & Q_2^*Q_2 \end{pmatrix}$$

より $Q_1^*Q_1 = I_N$ である．$D = \mathrm{diag}_{N,N}(\sigma_1, \cdots, \sigma_N)$ とおくと次が成り立つ．

$$\begin{aligned} A &= Q\begin{pmatrix} D \\ O \end{pmatrix}V^* = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}\begin{pmatrix} D \\ O \end{pmatrix}V^* \\ &= Q_1DV^* = Q_1V^*VDV^*. \end{aligned}$$

そこで $P = VDV^*$, $U = Q_1V^*$ とおけば，これが求める分解である．(1) は $A^*$

に (2) を適用すれば得られる．■

## 16.2　特異値標準形と一般逆行列

ここでは $A$ の特異値標準形を用いて一般逆行列の特徴づけを行う．以下では $A \in \boldsymbol{K}^{M\times N}$, $r = \mathrm{rank}(A)$ とする．また $\sigma_1 \geq \cdots \geq \sigma_r > 0$ を $A$ の 0 でない特異値とする．このとき定理 16.6 により，あるユニタリー行列 $U \in \boldsymbol{K}^{M\times M}$, $V \in \boldsymbol{K}^{N\times N}$ が存在し，

$$A = U \begin{pmatrix} \mathrm{diag}_{r,r}(\sigma_1, \cdots, \sigma_r) & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} V^*$$

と表せる (規約 11.16 参照)．次のことが成り立つ．

**定理 16.10**　$A, U, V, \sigma_1, \cdots, \sigma_r$ は本節のはじめに定めたものとする．

(1)　$B \in \boldsymbol{K}^{N\times M}$ が $A$ の一般逆行列であるための必要十分条件は，ある $C_{r,M-r} \in \boldsymbol{K}^{r\times(M-r)}$, $C_{N-r,r} \in \boldsymbol{K}^{(N-r)\times r}$, $C_{N-r,M-r} \in \boldsymbol{K}^{(N-r)\times(M-r)}$ が存在し，

$$B = V \begin{pmatrix} \mathrm{diag}_{r,r}(\sigma_1^{-1}, \cdots, \sigma_r^{-1}) & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} U^* \tag{16.4}$$

と表せることである．

(2)　$B$ が $A$ の反射型一般逆行列であるための必要十分条件は，ある $C_{r,M-r} \in \boldsymbol{K}^{r\times(M-r)}$, $C_{N-r,r} \in \boldsymbol{K}^{(N-r)\times r}$, $C_{N-r,M-r} \in \boldsymbol{K}^{(N-r)\times(M-r)}$ で

$$C_{N-r,M-r} = C_{N-r,r}\,\mathrm{diag}_{r,r}(\sigma_1, \cdots, \sigma_r)\, C_{r,M-r}$$

をみたすものが存在し，

$$B = V \begin{pmatrix} \mathrm{diag}_{r,r}(\sigma_1^{-1}, \cdots, \sigma_r^{-1}) & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} U^* \tag{16.5}$$

と表せることである．

(3)　$B$ が $A$ の最小 2 乗型一般逆行列であるための必要十分条件は，ある $C_{N-r,r} \in \boldsymbol{K}^{(N-r)\times r}$, $C_{N-r,M-r} \in \boldsymbol{K}^{(N-r)\times(M-r)}$ が存在し，

$$B = V \begin{pmatrix} \mathrm{diag}_{r,r}(\sigma_1^{-1}, \cdots, \sigma_r^{-1}) & O_{r\times(M-r)} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} U^* \tag{16.6}$$

と表せることである．

(4)　$B$ が $A$ のノルム最小型一般逆行列であるための必要十分条件は，ある $C_{r,M-r} \in \boldsymbol{K}^{r\times(M-r)}$, $C_{N-r,M-r} \in \boldsymbol{K}^{(N-r)\times(M-r)}$ が存在し，

$$B = V \begin{pmatrix} \mathrm{diag}_{r,r}\left(\sigma_1^{-1}, \cdots, \sigma_r^{-1}\right) & C_{r,M-r} \\ O_{(N-r)\times r} & C_{N-r,M-r} \end{pmatrix} U^* \tag{16.7}$$

と表されることである．

(5)　$B$ が $A$ のムーア-ペンローズ一般逆行列であるための必要十分条件は，

$$B = V \begin{pmatrix} \mathrm{diag}_{r,r}\left(\sigma_1^{-1}, \cdots, \sigma_r^{-1}\right) & O_{r\times(M-r)} \\ O_{(N-r)\times r} & O_{(N-r)\times(M-r)} \end{pmatrix} U^* \tag{16.8}$$

と表されることである．

**証明.** $C_{r,r} \in \boldsymbol{K}^{r\times r}$, $C_{r,M-r} \in \boldsymbol{K}^{r\times(M-r)}$, $C_{N-r,r} \in \boldsymbol{K}^{(N-r)\times r}$, $C_{N-r,M-r} \in \boldsymbol{K}^{(N-r)\times(M-r)}$ とし，

$$B = V \begin{pmatrix} C_{r,r} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} U^*$$

とおく．$D_{r,r} = \mathrm{diag}_{r,r}(\sigma_1, \cdots, \sigma_r)$ と表す．このとき $V^*V = I_N$, $U^*U = I_M$ より (以下，本定理の証明において，$\times$ は行列の積を表すものとする)

$$\begin{aligned} &ABA \\ &= U \begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} \begin{pmatrix} C_{r,r} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} \\ &\quad \times \begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} V^* \\ &= U \begin{pmatrix} D_{r,r}C_{r,r} & D_{r,r}C_{r,M-r} \\ O_{(M-r)\times r} & O_{(M-r)\times(M-r)} \end{pmatrix} \begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} V^* \\ &= U \begin{pmatrix} D_{r,r}C_{r,r}D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} V^*. \end{aligned}$$

(1)　以上の考察から，$ABA = A$ であるための必要十分条件は

$$D_{r,r}C_{r,r}D_{r,r} = D_{r,r}$$

である．すなわち

$$C_{r,r} = D_{r,r}^{-1}D_{r,r}D_{r,r}^{-1} = D_{r,r}^{-1} = \mathrm{diag}_{r,r}\left(\sigma_1^{-1}, \cdots, \sigma_r^{-1}\right)$$

である．

(2)　$B$ が $A$ の一般逆行列であるとする．このとき

$$
\begin{aligned}
&BAB\\
&= V\begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix}\begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix}\\
&\quad\times\begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix}U^*\\
&= V\begin{pmatrix} I_r & O_{r\times(N-r)} \\ C_{N-r,r}D_{r,r} & O_{(N-r)\times(N-r)} \end{pmatrix}\begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix}U^*\\
&= V\begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,r}D_{r,r}C_{r,M-r} \end{pmatrix}U^*.
\end{aligned}
$$

ゆえに，$BAB = B$ となるための必要十分条件は

$$C_{N-r,r}D_{r,r}C_{r,M-r} = C_{N-r,M-r}$$

である．逆にこの条件が成り立っていれば，(16.5) で表される行列 $B$ は一般逆行列であり，かつ反射型である．

(3)　$B$ が一般逆行列であるとする．このとき，

$$
\begin{aligned}
(AB)^* &= B^*A^*\\
&= U\begin{pmatrix} D_{r,r}^{-1} & C_{N-r,r}^* \\ C_{r,M-r}^* & C_{N-r,M-r}^* \end{pmatrix}V^*V\begin{pmatrix} D_{r,r} & O_{r\times(M-r)} \\ O_{(N-r)\times r} & O_{(N-r)\times(M-r)} \end{pmatrix}U^*\\
&= U\begin{pmatrix} D_{r,r}^{-1} & C_{N-r,r}^* \\ C_{r,M-r}^* & C_{N-r,M-r}^* \end{pmatrix}\begin{pmatrix} D_{r,r} & O_{r\times(M-r)} \\ O_{(N-r)\times r} & O_{(N-r)\times(M-r)} \end{pmatrix}U^*\\
&= U\begin{pmatrix} I_r & O_{r\times(M-r)} \\ C_{r,M-r}^*D_{r,r} & O_{(M-r)\times(M-r)} \end{pmatrix}U^*.
\end{aligned}
$$

一方，

$$
\begin{aligned}
AB &= U\begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix}\begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix}U^*\\
&= U\begin{pmatrix} I_r & D_{r,r}C_{r,M-r} \\ O_{(M-r)\times r} & O_{(M-r)\times(M-r)} \end{pmatrix}U^*
\end{aligned}
$$

であるから，$(AB)^* = AB$ であるための必要十分条件は，$C_{r,M-r} = O_{r\times(M-r)}$

である．逆に $B$ が (16.6) の形をしていれば，これまで示したことにより明らかに最小 2 乗型一般逆行列になっている．

(4)　$B$ を一般逆行列とし，

$$B = V \begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} U^*$$

とする．このとき

$$\begin{aligned}(BA)^* &= A^* B^* \\ &= V \begin{pmatrix} D_{r,r} & O_{r\times(M-r)} \\ O_{(N-r)\times r} & O_{(N-r)\times(M-r)} \end{pmatrix} \begin{pmatrix} D_{r,r}^{-1} & C_{N-r,r}^* \\ C_{r,M-r}^* & C_{N-r,M-r}^* \end{pmatrix} V^* \\ &= V \begin{pmatrix} I_r & D_{r,r} C_{N-r,r}^* \\ O_{(N-r)\times r} & O_{(N-r)\times(N-r)} \end{pmatrix} V^*.\end{aligned}$$

一方，

$$\begin{aligned}BA &= V \begin{pmatrix} D_{r,r}^{-1} & C_{r,M-r} \\ C_{N-r,r} & C_{N-r,M-r} \end{pmatrix} \begin{pmatrix} D_{r,r} & O_{r\times(N-r)} \\ O_{(M-r)\times r} & O_{(M-r)\times(N-r)} \end{pmatrix} V^* \\ &= V \begin{pmatrix} I_r & O_{r\times(N-r)} \\ C_{N-r,r} D_{r,r} & O_{(N-r)\times(N-r)} \end{pmatrix} V^*\end{aligned}$$

であるから，$(BA)^* = BA$ であるための必要十分条件は $C_{N-r,r} = O_{(N-r)\times r}$ である．逆に (16.7) で表される行列は $A$ のノルム最小型一般逆行列であることは容易に示せる．

(5)　$B$ が $A$ のムーア-ペンローズ一般逆行列であるとする．このとき $B$ は反射型，最小 2 乗型，ノルム最小 2 乗型であるから，

$$B = V \begin{pmatrix} D_{r,r}^{-1} & O_{r\times(M-r)} \\ O_{(N-r)\times r} & C_{N-r,M-r} \end{pmatrix} U^*$$

かつ

$$C_{N-r,M-r} = O_{(N-r)\times r}\, \mathrm{diag}_{r,r}(\sigma_1, \cdots, \sigma_r)\, O_{r\times(M-r)} = O_{(N-r)\times(M-r)}$$

でなければならない．逆に $B$ が (16.8) をみたしていれば，ムーア-ペンローズ一般逆行列の条件をみたしていることは容易にわかる．■

## 16.3　特異値分解と最小 2 乗解

$A$ を $(M,N)$ 行列とする．与えられた $\mathbf{b} \in \boldsymbol{K}^M$ に対して，本節では特異値分解を用いて，$A\mathbf{x} = \mathbf{b}$ の最小 2 乗解を構成する．一般に最小 2 乗解は一意的とは限らない．特異値分解を用いると，最小 2 乗解をすべて求めることができる．また特に $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解 $\mathbf{x}_{LS}$ で，

$$\|\mathbf{x}_{LS}\| = \min\left\{\|\mathbf{x}'\| : \mathbf{x}' \in \boldsymbol{K}^N \text{は } A\mathbf{x} = \mathbf{b} \text{ の最小 2 乗解}\right\} \tag{16.9}$$

をみたすものも求めることができる．

(16.9) をみたす $A\mathbf{x} = \mathbf{b}$ の最小 2 乗解を**ノルム最小の最小 2 乗解**という．ノルム最小の最小 2 乗解はただ一つしか存在しない (問題 16.3).

**定理 16. 11**　$A$ を階数 $r$ の $(M,N)$ 行列とする．$\{\mathbf{u}_1, \cdots, \mathbf{u}_M\}$ を $\boldsymbol{K}^M$ の正規直交基底，また $\{\mathbf{v}_1, \cdots, \mathbf{v}_N\}$ を $\boldsymbol{K}^N$ の正規直交基底で，$A$ の特異値分解 $A = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*$ を与えるものとする．$\mathbf{b} \in \boldsymbol{R}^M$ とする．このとき，

$$A\mathbf{x} = \mathbf{b}$$

の最小 2 乗解 $\mathbf{x}'$ はすべて

$$\mathbf{x}' = \sum_{i=1}^{r} \frac{1}{\sigma_i} \langle \mathbf{b}, \mathbf{u}_i \rangle \mathbf{v}_i + \sum_{i=r+1}^{N} \gamma_i \mathbf{v}_i,$$
$$\gamma_{r+1}, \cdots, \gamma_N \text{ は任意のスカラー}$$

によって与えられる．特に

$$\mathbf{x}_{LS} = \sum_{i=1}^{r} \frac{1}{\sigma_i} \langle \mathbf{b}, \mathbf{u}_i \rangle \mathbf{v}_i$$

とおくと，これは $A\mathbf{x} = \mathbf{b}$ のノルム最小の最小 2 乗解であり，$\mathbf{x}_{LS} = A^\dagger \mathbf{b}$ となっている．また最小 2 乗誤差は

$$\min_{\mathbf{x} \in \boldsymbol{K}^N} \|A\mathbf{x} - \mathbf{b}\| = \sqrt{\sum_{i=r+1}^{M} |\langle \mathbf{b}, \mathbf{u}_i \rangle|^2}$$

となっている．

**証明.** $\mathbf{x}$ を $\boldsymbol{K}^N$ の任意のベクトルとする．$\mathbf{x}$ はあるスカラー $\gamma_1, \cdots, \gamma_N$ により，$\mathbf{x} = \sum_{i=1}^{N} \gamma_i \mathbf{v}_i$ と表せる．また $A = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*$ であるから，

$$
\begin{aligned}
A\mathbf{x} &= \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^* \sum_{i=1}^{N} \gamma_i \mathbf{v}_i = \sum_{j=1}^{r} \sum_{i=1}^{N} \sigma_j \gamma_i \mathbf{u}_j \mathbf{v}_j^* \mathbf{v}_i \\
&= \sum_{j=1}^{r} \sum_{i=1}^{N} \sigma_j \gamma_i \mathbf{u}_j \delta_{ij} = \sum_{j=1}^{r} \sigma_j \gamma_j \mathbf{u}_j.
\end{aligned}
$$

$\mathbf{b} = \sum_{i=1}^{M} \langle \mathbf{b}, \mathbf{u}_i \rangle \mathbf{u}_i$ と表せるので

$$
A\mathbf{x} - \mathbf{b} = \sum_{j=1}^{r} (\sigma_j \gamma_j - \langle \mathbf{b}, \mathbf{u}_j \rangle) \mathbf{u}_j + \sum_{i=r+1}^{M} \langle \mathbf{b}, \mathbf{u}_i \rangle \mathbf{u}_i
$$

となり，

$$
\|A\mathbf{x} - \mathbf{b}\|^2 = \sum_{j=1}^{r} |\sigma_j \gamma_j - \langle \mathbf{b}, \mathbf{u}_j \rangle|^2 + \sum_{i=r+1}^{M} |\langle \mathbf{b}, \mathbf{u}_i \rangle|^2
$$

が得られる．ここで，$\sum_{i=r+1}^{M} |\langle \mathbf{b}, \mathbf{u}_i \rangle|^2$ は $\mathbf{x}$ に関係ない値である．これより，$\|A\mathbf{x} - \mathbf{b}\|$ が最小になるための $\mathbf{x}$ に関する必要十分条件は

$$
\sum_{j=1}^{r} |\sigma_j \gamma_j - \langle \mathbf{b}, \mathbf{u}_j \rangle|^2 = 0
$$

となること，すなわち

$$
\gamma_j = \frac{1}{\sigma_j} \langle \mathbf{b}, \mathbf{u}_j \rangle, \quad j = 1, \cdots, r
$$

となることである．このとき明らかに

$$
\|\mathbf{x}\|^2 = \|\mathbf{x}_{LS}\|^2 + \sum_{j=r+1}^{N} |\gamma_j|^2 \geq \|\mathbf{x}_{LS}\|^2
$$

である．定理 16.10 (5) により容易に $\mathbf{x}_{LS} = A^\dagger \mathbf{b}$ であることが示せる．■

**例 16.1**　方程式

$$
\begin{pmatrix} 1 & 4 \\ 2 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
$$

のノルム最小の最小 2 乗解を求めよ．

**答．** $A = \begin{pmatrix} 1 & 4 \\ 2 & 8 \end{pmatrix}$ とすると，

$$
A^* A = \begin{pmatrix} 5 & 20 \\ 20 & 80 \end{pmatrix}
$$

であり，その固有値は 85 と 0 である．またそれぞれその長さ 1 の固有ベクトルは，

$$\mathbf{v}_1 = \frac{1}{\sqrt{17}}\begin{pmatrix}1\\4\end{pmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{17}}\begin{pmatrix}-4\\1\end{pmatrix}$$

である．$\sigma_1 = \sqrt{85}$, $\sigma_2 = 0$ が特異値であり，

$$\begin{aligned}\mathbf{u}_1 &= \frac{1}{\sqrt{85}}\begin{pmatrix}1 & 4\\2 & 8\end{pmatrix}\frac{1}{\sqrt{17}}\begin{pmatrix}1\\4\end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix},\\ \mathbf{x}' &= \frac{1}{\sigma_1}\langle \mathbf{b}, \mathbf{u}_1\rangle \mathbf{v}_1\\ &= \frac{1}{\sqrt{85}}\left\langle \begin{pmatrix}0\\1\end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix}\right\rangle \frac{1}{\sqrt{17}}\begin{pmatrix}1\\4\end{pmatrix} = \frac{1}{85}\begin{pmatrix}2\\8\end{pmatrix}.\end{aligned}$$

■

**問題 16.3**　(1)　$X \subset \boldsymbol{K}^N$ が凸集合であるとは，任意の $\mathbf{x}, \mathbf{x}' \in X$, $0 \leq t \leq 1$ に対して，$t\mathbf{x} + (1-t)\mathbf{x}' \in X$ となることである．$A$ を $(M, N)$ 行列とするとき，

$$W = \left\{\mathbf{x} : \mathbf{x} \in \boldsymbol{K}^N \text{ は } A\mathbf{x} = \mathbf{b} \text{ の最小 2 乗解}\right\}$$

が $\boldsymbol{K}^N$ の凸集合であることを示せ．

(2)　$A\mathbf{x} = \mathbf{b}$ のノルム最小の最小 2 乗解はただ一つしかないことを示せ．

## 16.4　低階数の行列による近似とディジタル画像

本節ではまず，二つの行列がどのくらい違うかを測る尺度として，作用素ノルムと呼ばれる量を定義する．次にその尺度を用いて，$A \in \boldsymbol{K}^{M\times N}$ が与えられたとき，それより低い階数の行列 $B$ で，$A$ を近似することを考える．特異値分解を用いると，作用素ノルムで測って最も良い近似行列を求めることができる．最後に，ディジタル画像を行列と考えた場合に，この方法によって，ディジタル画像がより階数の少ない行列 (=画像) でどのように近似されるかを例示する．

**定義 16.4**　$(M, N)$ 行列 $A$ に対して，

$$\|A\|_{op} = \sup_{\mathbf{x}\in \boldsymbol{K}^N,\, \mathbf{x}\neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$$

とする．これを $A$ の**作用素ノルム**という．

はじめに $\|A\|_{op} < \infty$ であること，および $\|A - B\|_{op}$ の大きさと $A - B$ の各

成分 $a_{mn}-b_{mn}$ の大きさとの間の不等式を証明する．

**命題 16. 12** $(M,N)$ 行列 $A=(a_{mn})$ に対して，

$$\max_{m,n}|a_{mn}| \le \|A\|_{op} \le \sqrt{MN}\max_{m,n}|a_{mn}| < \infty$$

である．したがって $(M,N)$ 行列 $A=(a_{mn})$, $B=(b_{mn})$ に対して

$$\max_{m,n}|a_{mn}-b_{mn}| \le \|A-B\|_{op} \le \sqrt{MN}\max_{m,n}|a_{mn}-b_{mn}|$$

である．

**証明．** $a=\max_{m,n}|a_{mn}|$ とする．任意の $\mathbf{x}\in\boldsymbol{K}^N$ に対して，シュヴァルツの不等式より

$$\begin{aligned}\|A\mathbf{x}\|^2 &= \sum_{m=1}^{M}\left|\sum_{n=1}^{N}a_{mn}x_n\right|^2 \le \sum_{m=1}^{M}\sum_{n=1}^{N}|a_{mn}|^2\sum_{n=1}^{N}|x_n|^2 \\ &\le \sum_{m=1}^{M}\sum_{n=1}^{N}a^2\|\mathbf{x}\|^2 = MNa^2\|\mathbf{x}\|^2.\end{aligned}$$

ゆえに $\|A\|_{op} \le \sqrt{MN}a$ が成り立つ．いま $a=a_{jk}$ となる $a_{jk}$ をとる．このとき，

$$\begin{aligned}a=|a_{jk}| &\le \left(\sum_{m=1}^{M}|a_{mk}|^2\right)^{\frac{1}{2}} = \left\|A\mathbf{e}_k^{(N)}\right\| = \frac{\left\|A\mathbf{e}_k^{(N)}\right\|}{\left\|\mathbf{e}_k^{(N)}\right\|} \\ &\le \|A\|_{op}\end{aligned}$$

である．後半の主張は前半より明らかである． ■

**命題 16. 13** $A$ を $(M,N)$ 行列とする．任意の $\mathbf{x}\in\boldsymbol{K}^N$ に対して，

$$\|A\mathbf{x}\| \le \|A\|_{op}\|\mathbf{x}\|$$

である．

**証明．** $\mathbf{x}=\mathbf{0}$ の場合，命題の不等式は $0=0$ として成り立つ．

$\mathbf{x}\neq\mathbf{0}$ の場合，$\|A\|_{op}$ の定義から

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \le \|A\|_{op}$$

となり，命題が証明された． ■

**定理 16. 14** $A,B\in\boldsymbol{K}^{M\times N}$, $\alpha\in\boldsymbol{K}$ に対して

$$\|A+B\|_{op} \leq \|A\|_{op} + \|B\|_{op},$$
$$\|\alpha A\|_{op} = |\alpha|\,\|A\|_{op}.$$

また $\|A\|_{op} = 0$ であるための必要十分条件は $A = O_{M\times N}$ となることである．

**証明.** 命題 16.12 より，$\|A\|_{op} = 0 \Leftrightarrow A = O$ である．また $\alpha \in \boldsymbol{K}$ に対して

$$\|\alpha A\|_{op} = \sup_{\mathbf{x}\in \boldsymbol{K}^N,\, \mathbf{x}\neq \mathbf{0}} \frac{\|\alpha A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha| \sup_{\mathbf{x}\in \boldsymbol{K}^N,\, \mathbf{x}\neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha|\,\|A\|_{op}$$

である．さらに

$$\|(A+B)\,\mathbf{x}\| \leq \|A\mathbf{x}\| + \|B\mathbf{x}\| \leq \|A\|_{op}\,\|\mathbf{x}\| + \|B\|_{op}\,\|\mathbf{x}\|$$

より，

$$\|A+B\|_{op} \leq \|A\|_{op} + \|B\|_{op}$$

も得られる．■

$\|\cdot\|_{op}$ は，次の性質もみたす．

**定理 16. 15**　$A \in \boldsymbol{K}^{N_1\times N_2}$, $B \in \boldsymbol{K}^{N_2\times N_3}$ のとき，

$$\|AB\|_{op} \leq \|A\|_{op}\,\|B\|_{op}.$$

**証明.** $\|AB\mathbf{x}\| \leq \|A\|_{op}\,\|B\mathbf{x}\| \leq \|A\|_{op}\,\|B\|_{op}\,\|\mathbf{x}\|$ より容易に証明できる．■

$\|\cdot\|_{op}$ はユニタリー行列による作用に対して不変である．すなわち次のことが成り立つ．

**定理 16. 16**　$p, M, N$ を自然数とし，$N \leq M$ をみたすとする．

(1)　$U = (\mathbf{u}_1, \cdots, \mathbf{u}_N) \in \boldsymbol{K}^{M\times N}$ とし，$\mathbf{u}_1, \cdots, \mathbf{u}_N$ が $\boldsymbol{K}^M$ の正規直交系であるとする．このとき，$\|U\|_{op} = 1$ である．

(2)　$A \in \boldsymbol{K}^{N\times p}$ とする．$U \in \boldsymbol{K}^{M\times N}$ が (1) の条件をみたすとき，

$$\|UA\|_{op} = \|A\|_{op}$$

である．

(3)　$A \in \boldsymbol{K}^{N\times p}$ とする．$V \in \boldsymbol{K}^{p\times p}$ がユニタリー行列であれば，

$$\|AV\|_{op} = \|A\|_{op}$$

である．

**証明.** (1)　$\mathbf{x} = (x_j) \in \boldsymbol{K}^N$ に対して,

$$\|U\mathbf{x}\|^2 = \left\| \sum_{j=1}^{N} x_j \mathbf{u}_j \right\|^2 = \sum_{j=1}^{N} |x_j|^2 = \|\mathbf{x}\|^2 \tag{16.10}$$

であることによる.

(2)　$\mathbf{x} \in \boldsymbol{K}^p$ に対して, (16.10) より $\|UA\mathbf{x}\| = \|A\mathbf{x}\|$ であるから, 明らかに, $\|UA\|_{op} = \|A\|_{op}$ が成り立つ.

(3)　(1) より $\|V\|_{op} = 1$ である. したがって $\|AV\|_{op} \leq \|A\|_{op} \|V\|_{op} = \|A\|_{op}$ である. 一方, 任意の $\mathbf{x} \in \boldsymbol{K}^p$ に対して $\mathbf{y} = V^*\mathbf{x}$ とすると, $V$ はユニタリー行列であるから, $\mathbf{x} = V\mathbf{y}$ であり,

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|AV\mathbf{y}\|}{\|V\mathbf{y}\|} = \frac{\|AV\mathbf{y}\|}{\|\mathbf{y}\|} \leq \|AV\|_{op}$$

となる. ゆえに $\|A\|_{op} \leq \|AV\|_{op}$ も成り立つ. ■

$\|\cdot\|_{op}$ は次のように言い換えることもできる.

**定理 16.17**　$(M, N)$ 行列 $A$ に対して,

$$\|A\|_{op} = \sup\left\{\|A\mathbf{x}\| : \mathbf{x} \in \boldsymbol{K}^N,\ \|\mathbf{x}\| = 1\right\}$$

**証明.** $\|A\|_{2,2} = \sup\left\{\|A\mathbf{x}\| : \mathbf{x} \in \boldsymbol{K}^N,\ \|\mathbf{x}\| = 1\right\}$ とおく. $\|A\mathbf{x}\| \leq \|A\|_{op} \|\mathbf{x}\|$ より $\|A\|_{2,2} \leq \|A\|_{op}$ である. 一方, 任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{x} \neq \mathbf{0}$ に対して

$$\left\| \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| = 1$$

であり,

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \left\| A \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| \leq \|A\|_{2,2}$$

であるから, $\|A\|_{op} \leq \|A\|_{2,2}$ も得られる. ■

じつはこの作用素ノルムの値はちょうど最大の特異値と一致している.

**定理 16.18**　$A \in \boldsymbol{K}^{M \times N}$ とし, $\sigma_1 \geq \cdots \geq \sigma_{\min\{M,N\}}$ を $A$ の特異値とする. このとき, $\|A\|_{op} = \sigma_1$ である.

**証明.** $L = \min\{M, N\}$ とする. $A\mathbf{x} = \sum_{j=1}^{L} \sigma_j \mathbf{u}_j \mathbf{v}_j^*$ を $A$ の特異値分解とする. $A\mathbf{v}_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* \mathbf{v}_1 = \sigma_1 \mathbf{u}_1$ より, $\|A\mathbf{v}_1\| = \sigma_1$ であるから, $\|A\|_{op} \geq \sigma_1$ である. 一方, 任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\|\mathbf{x}\| = 1$ に対して

$$\|A\mathbf{x}\|^2 = \left\|\sum_{j=1}^{L} \sigma_j \mathbf{u}_j \mathbf{v}_j^* \mathbf{x}\right\|^2 = \left\|\sum_{j=1}^{L} \sigma_j \langle \mathbf{x}, \mathbf{v}_j \rangle \mathbf{u}_j\right\|^2$$
$$= \sum_{j=1}^{L} \sigma_j^2 |\langle \mathbf{x}, \mathbf{v}_j \rangle|^2 \leq \sigma_1^2 \sum_{j=1}^{L} |\langle \mathbf{x}, \mathbf{v}_j \rangle|^2 \leq \sigma_1^2 \|\mathbf{x}\|^2 = \sigma_1^2.$$

ゆえに $\|A\|_{op} \leq \sigma_1$ でもある．■

$A$ を階数 $r$ の $(M, N)$ 行列とする．$1 \leq k < r$ のとき，$A$ を階数 $k$ の行列で近似するという問題を考えてみよう．ここでは作用素ノルムを用いた近似，すなわち階数 $k$ の行列 $C$ で，$\|A - C\|_{op}$ が最も小さくなるようなものを求める．これについては，特異値分解を用いて，次のように最もよい近似を与える行列を構成することができる．

**定理 16. 19**　$A$ を階数 $r$ の $(M, N)$ 行列とする．$A = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*$ を $A$ の特異値分解とする．$1 \leq k < r$ に対して，

$$A_k = \sum_{j=1}^{k} \sigma_j \, \mathbf{u}_j \, \mathbf{v}_j^*$$

とおく．このとき $\mathrm{rank}(A_k) = k$ であり，

$$\min_{\substack{B \in \boldsymbol{K}^{M \times N} \\ \mathrm{rank}(B) \leq k}} \|A - B\|_{op} = \|A - A_k\|_{op} = \sigma_{k+1}$$

である．

定理の証明のため，次の補題を示す．

**補題 16. 20**　$L = \min\{M, N\}$ とする．このとき

$$\left\|\mathrm{diag}_{M,N}(\alpha_1, \cdots, \alpha_L)\right\|_{op} = \max\{|\alpha_1|, \cdots, |\alpha_L|\}$$

である．

**証明．** $\Delta = \mathrm{diag}_{M,N}(\alpha_1, \cdots, \alpha_L)$ とおく．$\max\{|\alpha_1|, \cdots, |\alpha_L|\} = |a_j|$ とする．$e^{i\theta_j}\alpha_j = |\alpha_j|$ なる $\theta_j \in [0, 2\pi)$ をとる．$\mathbf{x}^{(j)} = e^{i\theta_j}\mathbf{e}_j^{(N)}$ とおくと $\left\|\mathbf{x}^{(j)}\right\| = 1$ であり，$\Delta\,\mathbf{x}^{(j)} = |\alpha_j|\,\mathbf{e}_j^{(N)}$ であるから，$\left\|\Delta\mathbf{x}^{(j)}\right\| = |\alpha_j|$ である．ゆえに $\|\Delta\|_{op} \geq |\alpha_j|$ を得る．一方，任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\|\Delta\mathbf{x}\|^2 = \sum_{n=1}^{L} |\alpha_n x_n|^2 \leq |\alpha_j|^2 \sum_{n=1}^{L} |x_n|^2 \leq |\alpha_j|^2 \|\mathbf{x}\|^2$$

であるから，$\|\Delta\|_{op} \le |\alpha_j|$ も成り立つ． ■

**定理 16.19 の証明.** $U = (\mathbf{u}_1, \cdots, \mathbf{u}_M)$, $V = (\mathbf{v}_1, \cdots, \mathbf{v}_N)$ とすると，

$$\begin{aligned} U^* A_k V &= U^* \left( \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^* \right) V = \sum_{j=1}^{k} \sigma_j U^* \mathbf{u}_j \mathbf{v}_j^* V \\ &= \sum_{j=1}^{k} \sigma_j \begin{pmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_M^* \end{pmatrix} \mathbf{u}_j \mathbf{v}_j^* \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_N \end{pmatrix} \\ &= \sum_{j=1}^{k} \sigma_j \mathbf{e}_j^{(M)} \mathbf{e}_j^{(N)*} = \operatorname{diag}_{M,N}(\sigma_1, \cdots, \sigma_k, 0, \cdots, 0). \end{aligned}$$

ゆえに $\operatorname{rank}(A_k) = \operatorname{rank}(U^* A_k V) = k$ である．さらに定理 16.16，補題 16.20 を用いて

$$\begin{aligned} \|A - A_k\|_{op} &= \|U^*(A - A_k)V\|_{op} = \|U^* A V - U^* A_k V\|_{op} \\ &= \left\| \operatorname{diag}_{M,N}(0, \cdots, 0, \sigma_{k+1}, \cdots, \sigma_r, 0, \cdots, 0) \right\|_{op} = \sigma_{k+1}. \end{aligned}$$

定理を示すため，$M \times N$ 行列 $B$ で，$\operatorname{rank}(B) \le k$ をみたし $\|A - B\|_{op} < \sigma_{k+1}$ となるものが存在するとして矛盾を導く．仮定より $0 \ne \mathbf{x} \in \mathcal{N}(B)$ に対して，

$$\|A\mathbf{x}\| = \|(A - B)\mathbf{x}\| \le \|A - B\|_{op} \|\mathbf{x}\| < \sigma_{k+1} \|\mathbf{x}\| \tag{16.11}$$

である．一方，$\mathbf{v} \in \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} \subset \boldsymbol{K}^N$ とすると

$$\mathbf{v} = y_1 \mathbf{v}_1 + \cdots + y_{k+1} \mathbf{v}_{k+1} = V \begin{pmatrix} y_1 \\ \vdots \\ y_{k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (V\mathbf{y} \text{ とおく.})$$

したがって，$\|\mathbf{v}\| = \|V\mathbf{y}\| = \|\mathbf{y}\|$ である．$AV = U \operatorname{diag}_{M,N}(\sigma_1, \cdots, \sigma_L, 0 \cdots, 0)$ より

$$\begin{aligned} \|A\mathbf{v}\| &= \|AV\mathbf{y}\| = \left\| U \operatorname{diag}_{M,N}(\sigma_1, \cdots, \sigma_L, 0 \cdots, 0) \mathbf{y} \right\| \\ &= \left\| \operatorname{diag}_{M,N}(\sigma_1, \cdots, \sigma_L, 0 \cdots, 0) \mathbf{y} \right\| \end{aligned}$$

$$= \sqrt{\sum_{j=1}^{k+1} \sigma_j^2 y_j^2} \geq \sigma_{k+1} \|\mathbf{y}\| = \sigma_{k+1} \|\mathbf{v}\|. \tag{16.12}$$

ゆえに (16.11) と (16.12) より

$$\mathcal{N}(B) \cap \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} = \{\mathbf{0}\}$$

でなければならない．したがって $\mathcal{N}(B) \dotplus \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} \subset \boldsymbol{K}^N$ が得られ，

$$\dim \mathcal{N}(B) + \dim \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} \leq N$$

となる．ところで

$$k + \dim \mathcal{N}(B) \geq \operatorname{rank}(B) + \dim \mathcal{N}(B) = N$$

より $\dim \mathcal{N}(B) \geq N - k$ である．また $\dim \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} = k+1$ であるから，

$$\dim \operatorname{span}\{\mathbf{v}_1, \cdots, \mathbf{v}_{k+1}\} + \dim \mathcal{N}(B) \geq N + 1 > N.$$

よって矛盾が導かれた．■

$A$



**図 16.1**　500 × 500 ピクセル，256 階調のディジタル画像 $A$

定理 16.19 を画像処理に応用する．グレイ・スケールのディジタル画像は濃淡を表す数値の行列とみなすことができる．たとえば図 16.1 は 500 × 500 ピクセル，256 階調のディジタル画像であるが，これは濃淡を表す数値からなる (500, 500) 行列とみなすことができる．それを $A$ とおこう．

計算機による計算により，この画像の場合

$$\mathrm{rank}(A) = 500$$

である．このとき，$1 \leq k < 500$ に対して定理 16.19 で与えられる $A_k$ を考える．$\mathrm{rank}(A_k) = k$ である．$A_k$ がどのような画像になっているかを示しておこう．図 16.2 は $A_1$, $A_{10}$, $A_{20}$, $A_{50}$ を計算したものである．原画像は図 16.1 の $A$ であるが，階数 50 の画像 $A_{50}$ でも $A$ を視覚的によく近似していることがわかるであろう．

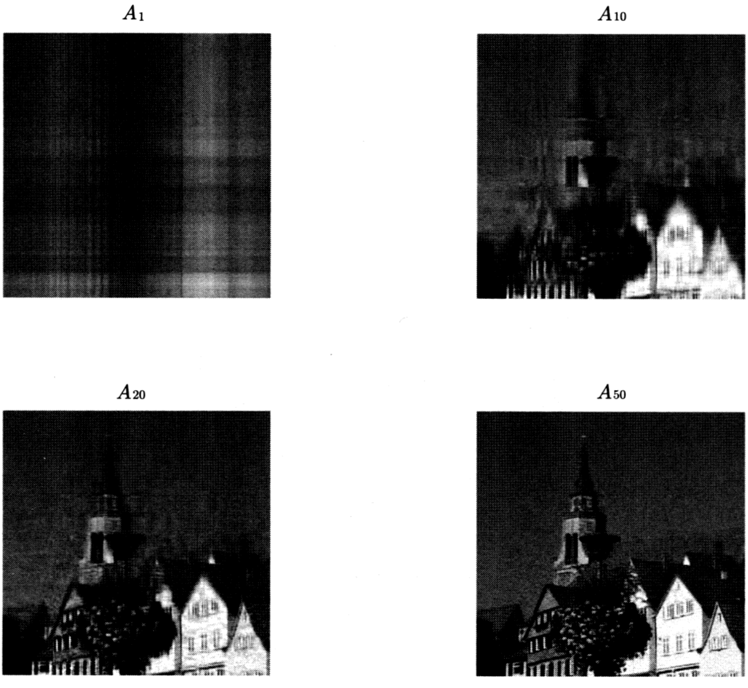



図 16.2

# 第17章　多変量解析と線形代数

調査や観測，あるいは実験などによって得られたデータから，その傾向や特性，データの背後に潜む要因などを分析する数学的な技法が多変量解析と呼ばれるものである．多変量解析の数学的な基盤の一つは線形代数にある．この章では，特に一般逆行列や特異値分解の多変量解析への応用について述べる．この章では $\boldsymbol{K} = \boldsymbol{R}$ とする．

**ポイント 17**　多変量解析はさまざまな場面で利用されている．線形代数の方法がどのように多変量解析に活かされているかを知ることは，さらに多変量解析を詳しく学ぶ際に役立つはずである．

## 17.1　いくつかの基本概念

多変量解析では，一般に調査あるいは観測したい対象を**個体**といい，調査項目を**変量**あるいは**変数**という．いくつかの個体に対して，必要な調査項目について調べた結果は，たとえば次のような表にまとめられる．

| | 変量 1 | 変量 2 | $\cdots$ | 変量 $p$ |
|---|---|---|---|---|
| 個体 1 | $x_{11}$ | $x_{12}$ | $\cdots$ | $x_{1p}$ |
| 個体 2 | $x_{21}$ | $x_{22}$ | $\cdots$ | $x_{2p}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| 個体 $N$ | $x_{N1}$ | $x_{N2}$ | $\cdots$ | $x_{Np}$ |

表 1

たとえば，あるクラスの学生たちの微分積分学，線形代数，統計学の期末試験

の成績を

| 学生番号 | 微分積分学 | 線形代数 | 統計学 |
|---|---|---|---|
| 1 | 78 | 82 | 80 |
| 2 | 55 | 70 | 68 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 100 | 40 | 50 | 20 |

表 2

のような表にまとめることができる．この場合はクラスの学生が個体であり，微分積分学，線形代数，統計学の試験結果が変量である．

さて，表 1 のデータは次のように行列で表すことができる．

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Np} \end{pmatrix} \tag{17.1}$$

このような行列は**データ行列**と呼ばれている．多変量解析はこのような行列を解析する技法を与えるものであり，そこで使われる数学的な道具はこれまで学んできた線形代数に依るところが大きい．

はじめに多変量解析でしばしば使われる用語の定義を述べる．以下では第 2.1 節で述べたコロン記法も用いる．この記法によれば (17.1) のデータ行列 $X$ に対して $X(j,k) = x_{jk}$ である．横ベクトル

$$X(j,:) = (x_{j1},\, x_{j2},\, \cdots,\, x_{jp})$$

は個体 $j$ に関するデータを表すベクトルであり，縦ベクトル

$$X(:,k) = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{Nk} \end{pmatrix}$$

は変量 $k$ に関するデータである．個体 $j$ に関するデータを表す横ベクトルは

$$\mathbf{x}_j = X(j,:)^T = \begin{pmatrix} x_{j1} \\ \vdots \\ x_{jp} \end{pmatrix} \in \boldsymbol{R}^p$$

のように縦ベクトルとして扱うこともある．

変量 $k$ に関するデータの平均値は多変量解析では基本的な役割を果たす量であり，それは

$$X_{\mathrm{ave}}(+,k) = \frac{1}{N}\sum_{j=1}^{N} X(j,k) = \frac{1}{N}\sum_{j=1}^{N} x_{jk}$$

と定義される．各変量の平均値からなるベクトル

$$X_{\mathrm{ave}} = \begin{pmatrix} X_{\mathrm{ave}}(+,1) \\ \vdots \\ X_{\mathrm{ave}}(+,p) \end{pmatrix} \in \boldsymbol{R}^p$$

を $X$ の**平均ベクトル**という．平均ベクトルは $\overline{\mathbf{x}}$ と表されることが多い．各変量のデータがその変量の平均値とどのくらいずれているかを表す行列

$$\begin{aligned}\widetilde{X} &= X - \mathbf{1}_N X_{\mathrm{ave}}^T \\ &= \begin{pmatrix} x_{11} - X_{\mathrm{ave}}(+,1) & \cdots & x_{1p} - X_{\mathrm{ave}}(+,p) \\ \vdots & \ddots & \vdots \\ x_{N1} - X_{\mathrm{ave}}(+,1) & \cdots & x_{Np} - X_{\mathrm{ave}}(+,p) \end{pmatrix} \in \boldsymbol{R}^{N\times p}\end{aligned}$$

を**平均偏差行列**という．$\widetilde{X}$ の $(j,k)$ 成分 $\widetilde{X}(j,k)$ は個体 $j$ の変量 $k$ に関する**平均偏差**(平均点からの差) と呼ばれている．さらに $(p,p)$ 行列

$$S_X = \frac{1}{N}\widetilde{X}^T\widetilde{X}$$

を**分散共分散行列**という．この行列は

$$S_X = \frac{1}{N}\begin{pmatrix} \widetilde{X}(1,1) & \cdots & \widetilde{X}(N,1) \\ \vdots & \ddots & \vdots \\ \widetilde{X}(1,p) & \cdots & \widetilde{X}(N,p) \end{pmatrix}\begin{pmatrix} \widetilde{X}(1,1) & \cdots & \widetilde{X}(1,p) \\ \vdots & \ddots & \vdots \\ \widetilde{X}(N,1) & \cdots & \widetilde{X}(N,p) \end{pmatrix}$$

$$= \frac{1}{N}\begin{pmatrix} \sum_{i=1}^{N}\widetilde{X}(i,1)\widetilde{X}(i,1) & \cdots & \sum_{i=1}^{N}\widetilde{X}(i,1)\widetilde{X}(i,p) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N}\widetilde{X}(i,p)\widetilde{X}(i,1) & \cdots & \sum_{i=1}^{N}\widetilde{X}(i,p)\widetilde{X}(i,p) \end{pmatrix}$$

と表せる．また $S_X$ の各成分は次のように表せる．

$$S_X(k_1,k_2) = \frac{1}{N}\sum_{i=1}^{N} \widetilde{X}(i,k_1)\,\widetilde{X}(i,k_2)$$
$$= \frac{1}{N}\sum_{i=1}^{N} \{x_{i,k_1} - X_{\mathrm{ave}}(+,k_1)\}\{x_{i,k_2} - X_{\mathrm{ave}}(+,k_2)\}$$
$$= \frac{1}{N}\left\langle \widetilde{X}(:,k_1), \widetilde{X}(:,k_2)\right\rangle, \quad (1 \le k_1, k_2 \le p).$$

このことから $S_X(k_1,k_2)$ はベクトル $\widetilde{X}(:,k_1)$, $\widetilde{X}(:,k_2)$ の内積となっている．$S_X(k_1,k_2)$ を変量 $k_1$, 変量 $k_2$ に関する**共分散**，また特に $S_X(k,k)$ を変量 $k$ の**分散**という．$\sqrt{S_X(k,k)}$ を変量 $k$ に関する**標準偏差**という．

**注意 17.1**　分散共分散行列は $\frac{1}{N}\widetilde{X}^T\widetilde{X}$ ではなく $\frac{1}{N-1}\widetilde{X}^T\widetilde{X}$ と定義されることも多い．本書では $\frac{1}{N}\widetilde{X}^T\widetilde{X}$ により定義する．

変量 $j$ に関するデータ $X(:,j) \in \boldsymbol{R}^N$ と変量 $k$ に関するデータ $X(:,k) \in \boldsymbol{R}^N$ に対して，$\widetilde{X}(:,j) \neq \mathbf{0}$, $\widetilde{X}(:,k) \neq \mathbf{0}$ の場合

$$r_X(j,k) = \frac{\left\langle \widetilde{X}(:,j), \widetilde{X}(:,k)\right\rangle}{\left\|\widetilde{X}(:,j)\right\| \left\|\widetilde{X}(:,k)\right\|} = \frac{S_X(j,k)}{\sqrt{S_X(j,j)}\sqrt{S_X(k,k)}}$$

とおき，これを**相関係数** (**ピアソンの積率相関係数**) という．幾何的な表現を用いれば，これは二つのベクトル $\widetilde{X}(:,j)$, $\widetilde{X}(:,k)$ のなす角の余弦である．定義より $|r_X(j,k)| \le 1$ であるが，特に $|r_X(j,k)| = 1$ の場合は，シュヴァルツの不等式の等号条件から，$\widetilde{X}(:,j)$, $\widetilde{X}(:,k)$ は線形従属である．また $r_X(j,k) = 0$ の場合は，二つのベクトルは直交している．相関係数が 0 に近いときは，二つのベクトル $\widetilde{X}(:,j)$, $\widetilde{X}(:,k)$ の相関は弱い，そして $|r_X(j,k)|$ が 1 に近いときは相関が強いという．また，$r_X(j,k) > 0$ ならば正の相関，$r_X(j,k) < 0$ ならば負の相関という．

$\widetilde{X}(:,j) \neq \mathbf{0}\,(j = 1, \cdots, p)$ の場合，

$$R_X = \begin{pmatrix} r_X(1,1) & \cdots & r_X(1,p) \\ \vdots & \ddots & \vdots \\ r_X(p,1) & \cdots & r_X(p,p) \end{pmatrix} \in \boldsymbol{R}^{p\times p}$$

とし，これを**相関係数行列**という．

$$D_X^{-\frac{1}{2}} = \mathrm{diag}_{p,p}\left(\frac{1}{\sqrt{S_X(1,1)}}, \cdots, \frac{1}{\sqrt{S_X(p,p)}}\right)$$

とおくと，分散共分散行列と相関係数行列の間には

$$R_X = D_X^{-\frac{1}{2}} S_X D_X^{-\frac{1}{2}} \tag{17.2}$$

の関係式が成り立つ (証明は各自試みよ).

**定理 17.1** (1)　$S_X, R_X$ は対称な非負定値行列である.

(2)

$$S_X = \frac{1}{N} X^T X - \overline{\mathbf{x}}\,\overline{\mathbf{x}}^T .$$

**証明.** (1)　分散共分散行列の定義と補題 16.1 より $S_X$ は対称な非負定値行列である．また (17.2) より

$$R_X = D_X^{-\frac{1}{2}} S_X D_X^{-\frac{1}{2}} = D_X^{-\frac{1}{2}} \widetilde{X}^T \widetilde{X} D_X^{-\frac{1}{2}} = \left(\widetilde{X} D_X^{-\frac{1}{2}}\right)^T \left(\widetilde{X} D_X^{-\frac{1}{2}}\right)$$

となっているから，補題 16.1 より $R_X$ に対しても (1) が示される.

(2) を示す.

$$\begin{aligned} S_X(k_1, k_2) &= \frac{1}{N} \sum_{i=1}^{N} x_{ik_1} x_{ik_2} - \frac{1}{N} \sum_{i=1}^{N} x_{ik_1} X_{\text{ave}}(+, k_2) \\ &\quad - \frac{1}{N} \sum_{i=1}^{N} x_{ik_2} X_{\text{ave}}(+, k_1) + X_{\text{ave}}(+, k_1) X_{\text{ave}}(+, k_2) \\ &= \frac{1}{N} \sum_{i=1}^{N} x_{ik_1} x_{ik_2} - X_{\text{ave}}(+, k_1) X_{\text{ave}}(+, k_2) \end{aligned}$$

である．一方,

$$\left(\frac{1}{N} X^T X\right)(k_1, k_2) = \frac{1}{N} \sum_{i=1}^{N} x_{ik_1} x_{ik_2},$$

$$\left(\overline{\mathbf{x}}\,\overline{\mathbf{x}}^T\right)(k_1, k_2) = X_{\text{ave}}(+, k_1) X_{\text{ave}}(+, k_2)$$

より定理が証明された. ■

この定理の証明中に現れた $\widetilde{X} D_X^{-\frac{1}{2}}$，すなわち

$$\widetilde{X} D_X^{-\frac{1}{2}} = \begin{pmatrix} \widetilde{X}(1,1) & \cdots & \widetilde{X}(1,p) \\ \vdots & \ddots & \vdots \\ \widetilde{X}(N,1) & \cdots & \widetilde{X}(N,p) \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{S_X(1,1)}} & & O \\ & \ddots & \\ O & & \frac{1}{\sqrt{S_X(p,p)}} \end{pmatrix}$$

$$
= \begin{pmatrix}
\frac{\widetilde{X}(1,1)}{\sqrt{S_X(1,1)}} & \frac{\widetilde{X}(1,2)}{\sqrt{S_X(2,2)}} & \cdots & \frac{\widetilde{X}(1,p)}{\sqrt{S_X(p,p)}} \\
\frac{\widetilde{X}(2,1)}{\sqrt{S_X(1,1)}} & \frac{\widetilde{X}(2,2)}{\sqrt{S_X(2,2)}} & \cdots & \frac{\widetilde{X}(2,p)}{\sqrt{S_X(p,p)}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\widetilde{X}(N,1)}{\sqrt{S_X(1,1)}} & \frac{\widetilde{X}(N,2)}{\sqrt{S_X(2,2)}} & \cdots & \frac{\widetilde{X}(N,p)}{\sqrt{S_X(p,p)}}
\end{pmatrix}
$$

を**標準得点行列**という．データ行列から標準得点行列を導くことを**データの標準化**という．標準化されたデータは平均が 0，分散が 1 になっている．すなわち $Z = \widetilde{X} D_X^{-\frac{1}{2}}$ とすると $k = 1, \cdots, p$ に対して

$$
\frac{1}{N} \sum_{j=1}^{N} Z(j,k) = 0, \quad S_Z(k,k) = 1 \tag{17.3}
$$

が成り立つ．

**問題 17.1**　(17.3) を証明せよ．

**問題 17.2**　$X \in \boldsymbol{R}^{N \times p}$ をデータ行列とし，$Z = \sum_{k=1}^{p} a_k X(:,k) \in \boldsymbol{R}^{N \times 1}$ とする．このとき，$S_Z$ を $S_X$ を用いて表せ．

## 17.2　回帰分析

$p+1$ 個の変量があり，ある変量 (目的変数) を残りの $p$ 個の変量 (説明変数) を用いて推定，説明あるいは近似することを**回帰分析**という．特に第 $p+1$ 変量を

$$
a_1 + a_2 \times \text{第}\,1\,\text{変量} + \cdots + a_{p+1} \times \text{第}\,p\,\text{変量}
$$

の形の一次関数を用いて近似をすることを**線形回帰**という．線形代数の用語を用いれば次のように言うことができる．あるデータ $\mathbf{y} = (y_j) \in \boldsymbol{R}^N$ とデータ行列 $X_0 \in \boldsymbol{R}^{N \times p}$ が与えられたとき，$X = (\mathbf{1}_N, X_0) \in \boldsymbol{R}^{N \times (p+1)}$ として，

$$
\|\mathbf{y} - X\mathbf{a}\| = \min_{\mathbf{b} \in \boldsymbol{R}^{p+1}} \|\mathbf{y} - X\mathbf{b}\| \tag{17.4}
$$

をみたす $\mathbf{a} = (a_j) \in \boldsymbol{R}^{p+1}$ を見出すことが線形回帰である．これは方程式 $\mathbf{y} = X\mathbf{b}$ の最小 2 乗解を求めることと同義であるかあら，第 15.3 章ですでに見たように，このような $\mathbf{a}$ として

$$\mathbf{a} = \left(X^T X\right)^- X^T \mathbf{y}$$

がある．もし $\mathrm{rank}(X) = p+1$ であれば，系 15.15 より，(17.4) をみたす $\mathbf{a}$ はただ一つ存在し，それは

$$\mathbf{a} = \left(X^T X\right)^{-1} X^T \mathbf{y}$$

により与えられる (この場合，$X^T X$ は正則行列であることに注意せよ)．このような $\mathbf{a}$ に対して，$\widehat{\mathbf{y}} = X\mathbf{a}$ を**予測値**，$\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}}$ を**残差**という．定義より，

$$\|\mathbf{e}\| = \min_{\mathbf{b} \in \boldsymbol{R}^{p+1}} \|\mathbf{y} - X\mathbf{b}\|$$

である．

次の等式が成り立つ．

**定理 17.2**　$\overline{y} = \dfrac{1}{N} \sum_{j=1}^{N} y_j$ とすると，

$$\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2 = \|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2 + \|\mathbf{e}\|^2$$

である．

**証明．** $\|\mathbf{y} - \widehat{\mathbf{y}}\| = E_{\mathcal{R}(X)}(\mathbf{y})$ であるから，$\widehat{\mathbf{y}} = P_{\mathcal{R}(X)}(\mathbf{y})$ である．ゆえに，$\mathbf{y} - \widehat{\mathbf{y}} \in \mathcal{R}(X)^\perp$ である．また，$\mathcal{R}(X) = \mathrm{span}\left\{\mathbf{1}_N, X\,(:,1)\,, \cdots, X\,(:,p)\right\}$ より，$\overline{y}\,\mathbf{1}_N \in \mathcal{R}(X)$ である．定義より $\widehat{\mathbf{y}} \in \mathcal{R}(X)$ でもあるから，$\langle \mathbf{y} - \widehat{\mathbf{y}}, \widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N \rangle = 0$ を得る．よって，ピタゴラスの定理より，

$$\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2 = \|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N - \widehat{\mathbf{y}} - \mathbf{y}\|^2 = \|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2 + \|\mathbf{e}\|^2$$

となる．■

**定義 17.1**　$\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2$ を**全平方和**，$\|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2$ を**回帰平方和**，$\|\mathbf{e}\|^2$ を**残差平方和**という．

$$R = \left( \frac{\|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2}{\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2} \right)^{\frac{1}{2}}$$

を**重相関係数**という．

定理 17.2 より，$R = 1$ ならば，$\mathbf{y} = \widehat{\mathbf{y}}$ である．重相関係数については，15.4 節も参照．

## 17.3 主成分分析

学生に対してある試験を行い，順位をつけたいとする．試験問題は全部で 5 問あったとする．このときどのように評価すべきか？　総点で評価してもよいが，総点が同点の者の順位をつけにくい．特に総点が同じでも解けた問題が異なる場合がある．そのようなときに試験結果を評価する一つの方法は，得点に適当な重みをつけて，その平均偏差の分散が最も大きくなるようにしたものによって評価することである (第 13.5 節参照)．その方法が本節で述べる主成分分析である．まず主成分の定義を述べる．次にその意味，そして適用例をあげる．

データ行列 $X = (x_{ij}) \in \boldsymbol{R}^{N\times p}$ $(p < N)$ に対して，その平均偏差行列

$$\widetilde{X} = \begin{pmatrix} x_{11} - X_{\mathrm{ave}}(+,1) & \cdots & x_{1p} - X_{\mathrm{ave}}(+,p) \\ \vdots & \ddots & \vdots \\ x_{N1} - X_{\mathrm{ave}}(+,1) & \cdots & x_{Np} - X_{\mathrm{ave}}(+,p) \end{pmatrix} \in \boldsymbol{R}^{N\times p}$$

の特異値分解をする．$\widetilde{X}$ の特異値を

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$$

とする．このとき特異値分解

$$\widetilde{X} = \sum_{j=1}^{p} \sigma_j \mathbf{u}_j \mathbf{v}_j^T$$

を与えるような $\boldsymbol{R}^N$ の正規直交基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ と $\boldsymbol{R}^p$ の正規直交基底 $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ が存在する．このような正規直交基底を分析の目的に合うように選んでとり，$\mathbf{v}_j$ を**第 $j$ 主成分ベクトル**といい，また $\langle \mathbf{x}_n - \overline{\mathbf{x}}, \mathbf{v}_j \rangle$ を個体 $n$ の**第 $j$ 主成分得点**という $(n = 1, \cdots, N)$．

さて，

$$U = (\mathbf{u}_1, \cdots, \mathbf{u}_N) \in \boldsymbol{R}^{N\times N},$$
$$V = (\mathbf{v}_1, \cdots, \mathbf{v}_p) \in \boldsymbol{R}^{p\times p}$$

は直交行列であり，

$$\widetilde{X} V = U \,\mathrm{diag}_{N,p}(\sigma_1, \cdots, \sigma_p) \tag{17.5}$$

をみたしている．定義より

$$\widetilde{X}V = \begin{pmatrix} (\mathbf{x}_1 - \overline{\mathbf{x}})^T \\ \vdots \\ (\mathbf{x}_N - \overline{\mathbf{x}})^T \end{pmatrix} (\mathbf{v}_1, \cdots, \mathbf{v}_p)$$

$$= \begin{pmatrix} \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_p \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_p \rangle \end{pmatrix}$$

であるから,

$$\begin{pmatrix} \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_p \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_p \rangle \end{pmatrix} = U \operatorname{diag}_{N,p}(\sigma_1, \cdots, \sigma_p)$$

$$= (\mathbf{u}_1, \cdots, \mathbf{u}_N) \begin{pmatrix} \sigma_1 & \cdots & O \\ & \ddots & \vdots \\ O & \cdots & \sigma_p \\ & O & \end{pmatrix}$$

$$= (\sigma_1 \mathbf{u}_1, \cdots, \sigma_p \mathbf{u}_p) \tag{17.6}$$

となっている.

これより

$$\left\langle \begin{pmatrix} \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_j \rangle \\ \vdots \\ \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_j \rangle \end{pmatrix}, \begin{pmatrix} \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_k \rangle \\ \vdots \\ \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_k \rangle \end{pmatrix} \right\rangle = \sigma_j \sigma_k \langle \mathbf{u}_j, \mathbf{u}_k \rangle = 0$$

が得られ，第 $j$ 主成分得点からなるベクトルと第 $k$ 主成分得点からなるベクトルは (零ベクトルでなければ) $j \neq k$ のとき直交しており，相関がないことがわかる.

**注意 17.2**　$S_X = \dfrac{1}{N}\widetilde{X}^T\widetilde{X}$ より，$S_X$ の固有値を

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$$

とおくと，$\widetilde{X}^T\widetilde{X}$ の固有値は $N\lambda_j$ であるから,

$$\sigma_j = \sqrt{N}\sqrt{\lambda_j}$$

である.

ここで主成分ベクトル $\mathbf{v}_1, \cdots, \mathbf{v}_p$ のもつ理論的な意味を考えてみよう．

1. $\mathbf{v}_1$ は $\mathbf{w} = (w_j)_{j=1,\cdots,p} \in \boldsymbol{R}^p$, $\|\mathbf{w}\| = 1$ なる重みをつけた合成得点

$$\begin{pmatrix} \sum_{j=1}^{p} (x_{1j} - X_{\text{ave}}(+,j))\, w_j \\ \vdots \\ \sum_{j=1}^{p} (x_{Nj} - X_{\text{ave}}(+,j))\, w_j \end{pmatrix} \tag{17.7}$$

の分散を最大にするような長さ 1 のベクトルである．

**証明.** (17.7) の分散は $\mathbf{w}^T S_X \mathbf{w}$ である．定理 13.26 と $\mathbf{v}_1$ が $S_X$ の最大固有値に属する長さ 1 の固有ベクトルであることから主張が示される． ■

2. $k = 2, \cdots, p$ とする．$\mathbf{v}_k$ は $\mathbf{v}_1, \cdots, \mathbf{v}_{k-1}$ と直交するような重み $\mathbf{w} = (w_j)_{j=1,\cdots,p} \in \boldsymbol{R}^p$, $\|\mathbf{w}\| = 1$ のうちで，合成得点

$$\begin{pmatrix} \sum_{j=1}^{p} (x_{1j} - X_{\text{ave}}(+,j))\, w_j \\ \vdots \\ \sum_{j=1}^{p} (x_{Nj} - X_{\text{ave}}(+,j))\, w_j \end{pmatrix}$$

の分散を最大にするような長さ 1 のベクトルである．

**証明.** クーラン-フィッシャーの定理 (定理 13.27) の証明を参考に容易に示せる．

■

**例題 17.1**　科目「数学」は線形代数と微分積分の講義からなっている．全体の 3 割に $A$ を与えることができる．この科目の履修者は 10 人で，線形代数と微分積分の試験を行ったところ次のような結果が得られた．ただし満点は 10 点である．

| 学生番号 | 線形 | 微積 | 総点 |
|---|---|---|---|
| 1 | 10 | 8 | 18 |
| 2 | 2 | 5 | 7 |
| 3 | 6 | 7 | 13 |
| 4 | 5 | 6 | 11 |
| 5 | 8 | 9 | 17 |
| 6 | 9 | 8 | 17 |
| 7 | 7 | 8 | 15 |
| 8 | 0 | 2 | 2 |
| 9 | 7 | 10 | 17 |
| 10 | 4 | 6 | 10 |

3 人の人に $A$ を与えるのであるが，総点から考えると，1 番の学生は $A$ であるが，5, 6, 9 番の学生は同点であるため，総点という尺度で測る限り順位がつけられない．そこでこのデータに関して第 1 主成分得点と第 2 主成分得点を求めてみる．平均ベクトルは

$$\overline{\mathbf{x}} = \frac{1}{10}\begin{pmatrix} 10+2+6+5+8+9+7+0+7+4 \\ 8+5+7+6+9+8+8+2+10+6 \end{pmatrix} = \begin{pmatrix} 5.8 \\ 6.9 \end{pmatrix}.$$

したがって平均偏差行列は

$$\widetilde{X} = \begin{pmatrix} 10-5.8 & 8-6.9 \\ 2-5.8 & 5-6.9 \\ 6-5.8 & 7-6.9 \\ 5-5.8 & 6-6.9 \\ 8-5.8 & 9-6.9 \\ 9-5.8 & 8-6.9 \\ 7-5.8 & 8-6.9 \\ 0-5.8 & 2-6.9 \\ 7-5.8 & 10-6.9 \\ 4-5.8 & 6-6.9 \end{pmatrix} = \begin{pmatrix} 4.2 & 1.1 \\ -3.8 & -1.9 \\ 0.2 & 0.1 \\ -0.8 & -0.9 \\ 2.2 & 2.1 \\ 3.2 & 1.1 \\ 1.2 & 1.1 \\ -5.8 & -4.9 \\ 1.2 & 3.1 \\ -1.8 & -0.9 \end{pmatrix}$$

である．$\widetilde{X}$ の特異値分解を計算機で計算すると

$$\widetilde{X} = U \begin{pmatrix} 11.254 & 0 \\ 0 & 2.8027 \\ O_{8\times 1} & O_{8\times 1} \end{pmatrix} \begin{pmatrix} 0.81934 & 0.57331 \\ 0.57331 & -0.81934 \end{pmatrix}$$

が得られる．したがって，次のような主成分得点が得られる．

| ＼ | 第 1 主成分得点 | 第 2 主成分得点 |
|---|---|---|
| 1 | 4.0719 | 1.5066 |
| 2 | −4.2028 | −0.62183 |
| 3 | 0.22120 | 0.032728 |
| 4 | −1.1715 | 0.27876 |
| 5 | 3.0065 | −0.45933 |
| 6 | 3.2525 | 0.93332 |
| 7 | 1.6138 | −0.2133 |
| 8 | −7.5614 | 0.68957 |
| 9 | 2.7605 | −1.8520 |
| 10 | −1.9908 | −0.29455 |

第 1 主成分得点は，

$$0.81934 \times [\text{線形代数の平均点との差}] + 0.57331 \times [\text{微分積分の平均点との差}]$$

で得られるので，総合評価を与えている．また第 2 主成分得点は

$$0.57331 \times [\text{線形代数の平均点との差}] - 0.81934 \times [\text{微分積分の平均点との差}]$$

であるので，第 2 主成分得点は，得点が正ならば線形代数が得意で，負ならば微分積分が得意ということになる．

学生番号 5 番と 6 番と 9 番の人は総合得点は同じであるが，分散を最大にした第 1 主成分得点を見ると，6 番，5 番，9 番の順になっている．6 番の人は平均点の低い線形代数で良い点を得たので第 1 主成分得点で高得点を獲得できたといえよう．

この例題では第 1 主成分得点により総合評価が得られ，さらに第 2 主成分まで見ることにより，より詳しい情報が得られている．一般の場合でも，第 1 主成分得点だけでなく，第 $m$ 主成分得点までを分析対象に入れた方がより細かい情報も考慮に入れているといえる．しかし一方で，できるだけ少ない情報でより良い評

価ができることが望ましいので，第 $m$ 主成分得点まで見るにしても，どのくらいまで見ればよいかがある程度わかると便利である．そのために寄与率という量が定義されている．これを定義するために，$k = 1, \cdots, p$ に対して第 $k$ 主成分得点からなるベクトルを

$$\mathbf{a}_k = \begin{pmatrix} \langle \mathbf{x}_1 - \overline{\mathbf{x}}, \mathbf{v}_k \rangle \\ \vdots \\ \langle \mathbf{x}_N - \overline{\mathbf{x}}, \mathbf{v}_k \rangle \end{pmatrix} \in \boldsymbol{R}^N$$

とおく．このベクトルのエネルギーが，すべての主成分得点のベクトルのエネルギーの総和の中で，どのくらいの比率を占めているかを考える．すなわち

$$C_k = \frac{\|\mathbf{a}_k\|^2}{\|\mathbf{a}_1\|^2 + \cdots + \|\mathbf{a}_p\|^2}$$

とする．(17.6) より

$$\|\mathbf{a}_k\|^2 = \|\sigma_k \mathbf{u}_k\|^2 = \sigma_k^2$$

であるから，

$$C_k = \frac{\sigma_k^2}{\sigma_1^2 + \cdots + \sigma_p^2}$$

となっている．$C_k$ を第 $k$ 主成分得点の**寄与率**といい，寄与率は $C_k$%という．また $C_1 + \cdots + C_k$ を第 $k$ 主成分得点までの**累積寄与率**という．

これまで述べた主成分分析は平均偏差行列 $\widetilde{X}$ の特異値分解を用いて行ったが，データの単位が不揃いの場合などはデータの標準化を行い，標準得点行列 $\widetilde{X} D_X^{-\frac{1}{2}}$ の特異値分解を用いて主成分分析を行う．この場合，分散共分散行列 $S_X$ の代わりの役は相関係数行列 $R_X$ が演ずる．

# 第18章　離散フーリエ解析への応用

本章では離散フーリエ変換について述べる．離散フーリエ変換はディジタル信号処理をはじめさまざまな分野で利用されている有用な線形写像の一つである．

本章ではこれまでの章とは記法を変えて，ベクトル $\mathbf{x}$ の成分を $x[k]$ と表し，番号 $k$ は 0 から付けることにする．すなわち，$\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$\mathbf{x} = \begin{pmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{pmatrix}$$

のように成分の番号をつける．その理由は，フーリエ解析では $x[0]$ の 0 を周波数 0 の成分に対応させると便利だからである．またこれに合わせて行列も

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1,0} & a_{N-1,1} & \cdots & a_{N-1,N-1} \end{pmatrix}$$

のように番号は 0 からつけるものとする．

以下では $\boldsymbol{K} = \boldsymbol{C}$ とし

$$\boldsymbol{Z}_N = \{0, 1, \cdots, N-1\} \tag{18.1}$$

と表す．また本章と次章では $i$ により純虚数を表す．すなわち

$$i = \sqrt{-1}$$

である．前章まで記号 $i$ は $a_i$, $i = 1, \cdots, n$ のように添え字として用いることもあったが，混乱を避けるため本章に限り $i$ を添え字として使うことはない．

**ポイント 18**　離散フーリエ変換は応用上極めて重要な変換である．本書では，離散フーリエ変換のもつ性質，ならびに離散フーリエ変換によって何がわかるかを理解してほしい．なお計算機では，高速フーリエ変換と呼ばれる計算量を大幅に軽減したアルゴリズムで計算されている．

## 18.1　フーリエ解析とは何か

離散フーリエ解析のもととなったものは，フランスの数学者 J. B. J. フーリエによって創始されたフーリエ解析である．フーリエは，周期 1 の関数 $f(x)$，すなわち $f(x+1)=f(x)\ (x \in \boldsymbol{R})$ をみたす関数はすべて

$$f(x)=\sum_{k=-\infty}^{\infty} c_k(f)e^{2\pi ikx}, \tag{18.2}$$

ただし，

$$c_k(f)=\int_0^1 f(t)e^{-2\pi ikt}dt \tag{18.3}$$

のように三角関数の和で表されると主張した (注意 18.1 参照)．ここで関数

$$e^{2\pi ikx}=\cos(2\pi kx)+i\sin(2\pi kx)$$

は複素正弦波と呼ばれているもので, $k$ をその複素正弦波の周波数という．周波数の絶対値が小さければ, 複素正弦波の実数部分 $\cos(2\pi kx)$ および虚数部分 $\sin(2\pi kx)$ はともに周期の大きな波を表し，周波数の絶対値が大きくなると周期の小さい波となっている．このことは，周期 1 のすべての関数が周波数 $0, 1, 2, \cdots$ の正弦波をたし合わせてできていて，さらに $c_k$ および $c_{-k}$ は周波数 $|k|$ の複素正弦波 $e^{2\pi ikx}$, $e^{-2\pi ikx}$ をどのくらいの量加えれば良いかを示している．言い換えれば，すべての周期 1 の関数は周波数 $0, 1, 2, \cdots$ の複素正弦波から合成できるともいえる．このように周期 1 の関数が (18.2), (18.3) のように表せるという主張を「フーリエの原理」と呼ぶ．

ところで，周期 1 の関数 $f(x)$ をフーリエの原理に基づいて分析するためには，無限個のデータ $\{f(x) : x \in [0,1)\}$ を処理しなければならない．しかし，われわれが実際に計算機などで扱えるデータは有限個のものである．そのため，実用的にはたとえば自然数 $N$ をとり，

$$t_n=\frac{n}{N}, \quad n=0,\cdots,N-1$$

を分点とする $[0,1]$ の分割 $0=t_0<\cdots<t_{N-1}<1$ でサンプルとしてとった有限個のデータ $\{f(t_n):n=0,\cdots,N-1\}$ を処理することになる．$N$ を**サンプル数**，$t_n$ を**サンプル点**という．このような有限個のデータを解析するために考えられたものが，フーリエ解析の有限な離散化である (有限) 離散フーリエ解析である．

**注意 18.1**　フーリエは周期 1 の関数に対して，フーリエの原理が成り立つと主張したが，それを厳密に証明したわけではなかった．そのため彼の主張に対して多くの反論が提起された．実際に (18.2) をある点で満たさないような連続関数も構成された．しかしながら，かなり多くの実用的な関数に対しては，フーリエの原理が成立することや，(18.2) の級数の収束を各点での収束ではなく，$L^2$ 収束などの意味に解釈すれば，連続関数をはじめ多くの関数に対してフーリエの原理が成立することが知られている．詳細はたとえば [ 新井 ] などフーリエ解析の教科書を参照してほしい．

**注意 18.2**　いくつかの用語を定義しておく．$\boldsymbol{R}$ または区間 $I$ 上で定義された実数値あるいは複素数値の関数 $x_c(t)$ を**連続時間信号**あるいは**アナログ信号**という．信号解析で基本的な連続時間信号は $f$ を実数として

$$z(t)=Ae^{i\theta}e^{2\pi ift},$$

あるいは，この信号の実部

$$x(t)=A\cos(2\pi ft+\theta) \tag{18.4}$$

である．ここで $A$ を**振幅**，$f$ を**周波数**，$\theta$ を**位相シフト**という．$t$ が時間 $[s]$ を表すとき，$f$ の単位は [Hz] である．また，

$$2\pi f$$

を**角周波数**という．連続時間信号に対して，整数全体のなす集合 $\boldsymbol{Z}$ のある部分集合 $J$ 上で定義された実数値あるいは複素数値の関数 $x[n]$ を**離散時間信号**という (ここでは扱わないが離散時間信号の各値 $x[n]$ が有限ビット長で表せるものを**ディジタル信号**という)．たとえば連続時間信号 $x_c(t)$ の値を $T_s$ 秒の間隔で採取することにより離散時間信号

$$x[n]=x_c(nT_s),\quad n\in\boldsymbol{J}$$

を作ることができる．このように連続時間信号から離散時間信号を作る操作を**サンプリング**といい，$T_s$ を**サンプリング周期**，あるいは**サンプリング間隔**，$f_s=\dfrac{1}{T_s}$

を**サンプリング周波数**という．

離散時間信号のなかで基本的なものの一つは

$$z[n] = Ae^{i\theta}e^{2\pi i\widehat{f}n},$$

あるいはその実部

$$x[n] = A\cos\left(2\pi\widehat{f}n + \theta\right)$$

である．ここで $A$ を振幅，$\widehat{f}$ を**正規化周波数**，$\theta$ を位相シフトという．たとえば $x[n]$ がサンプリング間隔 $T_s$ のサンプリングにより (18.4) から作った離散時間信号であるとき，

$$x[n] = A\cos(2\pi f T_s n + \theta) = A\cos\left(2\pi\widehat{f}n + \theta\right)$$

であり，周波数 $f$ と正規化周波数 $\widehat{f}$ には

$$\widehat{f} = fT_s = \frac{f}{f_s}$$

の関係がある．

## 18.2　フーリエ基底

本書では $e^{ix}$ を $\exp(ix)$ とも表す．$N$ をサンプル数とし，複素正弦波

$$\exp(2\pi ikx), \quad 0 \leq x < 1$$

のサンプル点

$$t_n = \frac{n}{N}, \quad n = 0, \cdots, N-1$$

における値

$$\exp\left(2\pi ik\,\frac{n}{N}\right)$$

を考える．これに $\dfrac{1}{\sqrt{N}}$ をかけたものを

$$F_k^{(N)}[n] = \frac{1}{\sqrt{N}}\exp\left(2\pi ik\,\frac{n}{N}\right), \quad n = 0, \cdots, N-1$$

とし，これを成分とするベクトル

$$\mathbf{F}_k^{(N)} = \begin{pmatrix} F_k^{(N)}[0] \\ \vdots \\ F_k^{(N)}[N-1] \end{pmatrix} \in \boldsymbol{K}^N$$

を定義する．図 18.1 は，$N=8$ の場合に $\mathbf{F}_k^{(8)}$ がどのようなベクトルであるかを図示したものである．

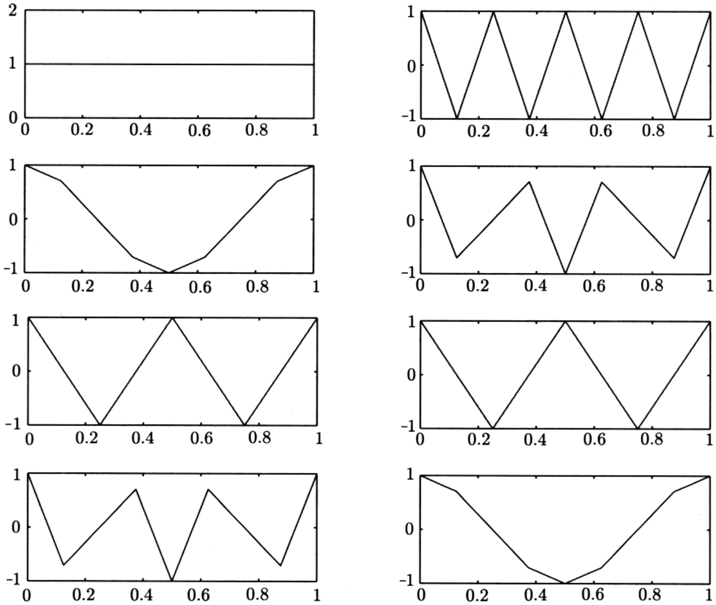


**図 18.1**

図は $(t_0, \mathrm{Re}\,\mathbf{F}_k^{(8)}[t_0]), (t_1, \mathrm{Re}\,\mathbf{F}_k^{(8)}[t_1]), \cdots, (t_N, \mathrm{Re}\,\mathbf{F}_k^{(8)}[t_N])$ をプロットし，プロットした隣同士の点を線分で結んだものである．左側は上から $k=0,1,2,3$, そして右側は上から $k=4,5,6,7$ のグラフを並べて記している．

いま

$$\omega\ (=\omega_N) = \exp\left(\frac{2\pi i}{N}\right)$$

とおくと，

$$F_k^{(N)}[n] = \frac{1}{\sqrt{N}}\omega^{kn}$$

と表される．

今後，サンプル数が固定されていて特に混乱が生じない場合，

$$F_k^{(N)}[n] = F_k[n], \quad \mathbf{F}_k^{(N)} = \mathbf{F}_k$$

と略記する．

次のことが成り立つ．

**定理 18.1**　$\{\mathbf{F}_0, \mathbf{F}_1, \cdots, \mathbf{F}_{N-1}\}$ は $\boldsymbol{K}^N$ の正規直交基底である．

**証明.** $j, k = 0, \cdots, N-1$ に対して

$$\langle \mathbf{F}_j, \mathbf{F}_k \rangle = \frac{1}{N} \sum_{n=0}^{N-1} \omega^{nj} \omega^{-nk} = \frac{1}{N} \sum_{n=0}^{N-1} \omega^{(j-k)n}$$

である．ここで等式 $(1-z)(1+z+\cdots+z^{N-1}) = 1-z^N$ を $z = \omega^{j-k}$ に対して用いると

$$\left(1-\omega^{j-k}\right) \sum_{n=0}^{N-1} \omega^{(j-k)n} = 1 - \omega^{(j-k)N} = 0$$

が得られる．もし $j \neq k$ ならば $\omega^{j-k} \neq 1$ であるから，

$$\sum_{n=0}^{N-1} \omega^{(j-k)n} = 0 \tag{18.5}$$

となり，$\langle \mathbf{F}_j, \mathbf{F}_k \rangle = 0$ である．また $j = k$ ならば明らかに $\langle \mathbf{F}_j, \mathbf{F}_k \rangle = 1$ である．

■

このことと定理 10.8 を組み合わせて，次の定理を得る．これは有限離散フーリエ解析におけるフーリエの原理である．

**定理 18.2**　任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x} = \sum_{k=0}^{N-1} \widehat{x}[k]\, \mathbf{F}_k \tag{18.6}$$

と表せる．ただしここで，

$$\widehat{x}[k] = \langle \mathbf{x}, \mathbf{F}_k \rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \exp\left(-2\pi i k \frac{n}{N}\right), \quad k = 0, \cdots, N-1$$

である．

**定義 18.1**　定理 18.2 において，$\widehat{x}[k]$ を $\mathbf{x}$ の**第** $k$ **フーリエ係数**という ($k = 0, \cdots, N-1$).

ここで定理 18.2 とフーリエ係数の意味を述べておこう．たとえば時刻 0 から $T$ まで音声信号が発せられていて，時刻 $0 = t_0 < t_1 < \cdots < t_{N-1} < T$ においてその空気圧を測定した数値を $x[0], \cdots, x[N-1]$ とする．このとき，点 $(t_j, x[j])$ をプロットしたグラフを描くと，たとえば図 1.1 のようになる (ただし添字の番号が 1 つずれている)．さて，サンプルをとった値を並べて $\mathbf{x}^T = (x[0], \cdots, x[N-1])$

によりベクトル $\mathbf{x} \in \boldsymbol{K}^N$ を定義する．そして，これに定理 18.2 を適用すると，(18.6) は音声信号を表す $\mathbf{x}$ が $\mathbf{F}_0, \cdots, \mathbf{F}_{N-1}$ の一次結合で表されることを意味している．そして，$\widehat{x}[k]$ は $\mathbf{F}_k$ がこの音声信号の中にどれだけ含まれているかを表している．逆に言えば，$\mathbf{F}_k$ を $\widehat{x}[k]$ 倍して $\sum_{k=0}^{N-1} \widehat{x}[k]\,\mathbf{F}_k$ と重ね合わせれば，もともとの音声信号を離散化した $\mathbf{x}$ が再現できることを意味している．

ここで注意すべき点がある．複素正弦波 $e^{2\pi ikx}$ の実数部分，虚数部分は，$|k|$ が大きくなればそれだけ周期の小さな波形を表している．しかし，図 18.1 をみればわかるように，$k$ が大きいからといって周期の小さな波を表しているとは限らない．たとえば $\mathbf{F}_7^{(8)} = \mathbf{F}_1^{(8)}$ である．一般に次のことが成り立っている．

**命題 18.3** $\overline{\mathbf{F}_{N-k}^{(N)}} = \mathbf{F}_k^{(N)}$. ((2.5) 参照)

このことから，$\mathbf{F}_k^{(N)}$ は $k$ が $\dfrac{N}{2}$ の近くにある場合は周波数が大きい高周波を表し，$k$ が 0 あるいは $N-1$ の近くにある場合は周波数の小さい低周波を表している．

次に離散フーリエ変換を定義する．

**定義 18.2** $\mathbf{x} \in \boldsymbol{K}^N$ のフーリエ係数を並べたベクトルを

$$\widehat{\mathbf{x}} = \begin{pmatrix} \widehat{x}[0] \\ \vdots \\ \widehat{x}[N-1] \end{pmatrix}$$

と表す．$\mathbf{x} \in \boldsymbol{K}^N$ に $\sqrt{N}\,\widehat{\mathbf{x}} \in \boldsymbol{K}^N$ を対応させる写像を $\mathcal{F}_N$ とおき，$N$ **点離散フーリエ変換**あるいは単に**離散フーリエ変換**という．

$$\mathcal{F}_N(\mathbf{x}) = \sqrt{N}\,\widehat{\mathbf{x}}$$

と表す．ただし，$N$ について混乱のない場合は $\mathcal{F}_N$ を単に $\mathcal{F}$ と略記する．

離散フーリエ変換では，$\widehat{\mathbf{x}}$ に $\sqrt{N}$ をかけていることに注意せよ．この場合，定義から，

$$\mathcal{F}_N(\mathbf{x})[k] = \sqrt{N}\,\widehat{x}[k] = \sum_{n=0}^{N-1} x[n] \exp\left(-2\pi ik\,\frac{n}{N}\right), \quad k = 0, \cdots, N-1$$

である．

$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$, $\alpha, \beta \in \boldsymbol{K}$ に対して，

$$\widehat{\alpha x + \beta y}[k] = \langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{F}_k \rangle = \alpha \langle \mathbf{x}, \mathbf{F}_k \rangle + \beta \langle \mathbf{y}, \mathbf{F}_k \rangle$$
$$= \alpha \widehat{x}[k] + \beta \widehat{y}[k]$$

であるから，次のことがわかる．

**定理 18.4**　$N$ 点離散フーリエ変換 $\mathcal{F}$ は $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像である．

$N$ 点離散フーリエ変換を表現する行列を求めておく．

$$\widehat{x}[k] = \sum_{n=0}^{N-1} x[n] \overline{F_k[n]} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \omega^{-kn}.$$

したがって，

$$\widehat{\mathbf{x}} = \frac{1}{\sqrt{N}} \begin{pmatrix} \sum_{n=0}^{N-1} x[n] \omega^{-0\cdot n} \\ \vdots \\ \sum_{n=0}^{N-1} x[n] \omega^{-(N-1)n} \end{pmatrix}$$

$$= \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \cdots & \omega^{-(N-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \cdots & \omega^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(N-1)} & \omega^{-(N-1)2} & \cdots & \omega^{-(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{pmatrix}$$

である．そこで

$$F\ (= F_N) = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \cdots & \omega^{-(N-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \cdots & \omega^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(N-1)} & \omega^{-(N-1)2} & \cdots & \omega^{-(N-1)(N-1)} \end{pmatrix}$$

とおき，これを $N$ **次フーリエ行列**という．次のことが成り立つ．

**定理 18.5**　$N$ 点離散フーリエ変換 $\mathcal{F}$ は $N$ 次フーリエ行列 $F$ を用いて次のように表せる．

$$\mathcal{F}(\mathbf{x}) = \sqrt{N} F \mathbf{x}, \quad \mathbf{x} \in \boldsymbol{K}^N.$$

フーリエ行列は次の性質をもつ．

**定理 18.6** $N$ 次フーリエ行列 $F$ はユニタリー行列である．

**証明.** $F^* = \begin{pmatrix} \mathbf{F}_0 & \cdots & \mathbf{F}_{N-1} \end{pmatrix}$ であり，$\{\mathbf{F}_0, \cdots, \mathbf{F}_{N-1}\}$ は $\boldsymbol{K}^N$ の正規直交基底であるから，$F^*$ はユニタリー行列である．ゆえに $F$ もユニタリー行列である．■

ユニタリー行列のもつ性質から次の系が導かれる．

**系 18.7（離散パーセヴァルの等式）** $\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\langle \mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y}) \rangle = N \langle \mathbf{x}, \mathbf{y} \rangle$$

である．特に

$$\|\mathcal{F}(\mathbf{x})\| = \sqrt{N} \|\mathbf{x}\|$$

が成り立つ．

**系 18.8** $N$ 点離散フーリエ変換 $\mathcal{F}$ は逆写像 $\mathcal{F}^{-1}$ をもち，

$$\mathcal{F}^{-1}(\mathbf{x}) = \frac{1}{\sqrt{N}} F^* \mathbf{x}$$

で与えられる．

フーリエ解析では，原信号 $\mathbf{x}$ を解析するために，$\mathcal{F}(\mathbf{x})$ を解析し，原信号にどのくらいの周波数の $\mathbf{F}_k$ が含まれているかを見ることにより，$\mathbf{x}$ の性質を研究する．この方法は調和解析とも呼ばれる．

## 18.3 フィルタリングとその応用(ノイズ除去)

信号解析では，ノイズ除去，画像データの鮮鋭化，エッジ強調などにフィルタリングという操作が行われることがある．フィルタリングはたたみ込み積 (convolution) という演算によって定義される．この節ではたたみ込み積の定義と基本的な性質を学ぶ．たたみ込み積は応用上重要な演算である．なお，ここで扱うたたみ込み積は循環たたみ込み積と呼ばれるものである．

本節では，次の規約を設ける．

**規約 18.9（ベクトルの周期拡張）** $\boldsymbol{K}^N$ のベクトル $\mathbf{x} = (x[n])_{n=0,\cdots,N-1}$ に

対して

$$x[j+kN]=x[j],\quad j=0,\cdots,N-1,\quad k=\pm1,\pm2,\cdots$$

と定める．このように定義した $(x[n])_{n\in\boldsymbol{Z}}$ を $\mathbf{x}=(x[n])_{n=0,\cdots,N-1}$ の周期拡張という．$\boldsymbol{K}^N$ のベクトル $\mathbf{x}=(x[n])_{n=0,\cdots,N-1}$ に対して，$(x[n])_{n\in\boldsymbol{Z}}$ は $\mathbf{x}$ の周期拡張を表すものとする．

循環たたみ込み積は次のように定義される．

**定義 18.3**　$\mathbf{x},\mathbf{y}\in\boldsymbol{K}^N$ に対して，

$$(x *_c y)[j]=\sum_{k=0}^{N-1}x[k]\,y[j-k],\quad j=0,\cdots,N-1$$

とし，

$$\mathbf{x} *_c \mathbf{y}=((x *_c y)[j])\in\boldsymbol{K}^N$$

を $\mathbf{x}$ と $\mathbf{y}$ の**循環たたみ込み**という．

フィルタリングはこの循環たたみ込みを用いて次のように定義される．

**定義 18.4**　$\mathbf{a}=(a[n])_{n=0,\cdots,N-1}\in\boldsymbol{K}^N$ とする．このとき

$$\mathcal{C}_{\mathbf{a}}(\mathbf{x})=\mathbf{a} *_c \mathbf{x},\quad \mathbf{x}\in\boldsymbol{K}^N$$

とし，$\mathbf{x}\in\boldsymbol{K}^N$ に $\mathcal{C}_{\mathbf{a}}(\mathbf{x})\in\boldsymbol{K}^N$ を対応させる写像を $\mathcal{C}_{\mathbf{a}}$ と表し，シンボル $\mathbf{a}$ の**たたみ込み作用素**という．$\mathbf{x}$ にたたみ込み作用素 $\mathcal{C}_{\mathbf{a}}$ を施すことをフィルタ $\mathbf{a}$ による**フィルタリング**という．

循環たたみ込み積の定義から次のことは明らかである．

**定理 18.10**　定義 18.4 で定義されたシンボル $\mathbf{a}$ のたたみ込み作用素は $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像である．

循環たたみ込み積の基本的な性質をいくつか証明する．まず循環たたみ込み積が可換な演算であることを示す．

**命題 18.11**　$\mathbf{x},\mathbf{y}\in\boldsymbol{K}^N$ に対して，

$$\mathbf{x} *_c \mathbf{y}=\mathbf{y} *_c \mathbf{x}$$

である．

**証明.** $k' = j - k$ と変換すると，規約 18.9 より

$$(x *_c y)[j] = \sum_{k=0}^{N-1} x[k]\, y[j-k] = \sum_{k'=j}^{j-(N-1)} x[j-k']\, y[k']$$

$$= \sum_{k'=0}^{N-1} x[j-k']\, y[k'] = (y *_c x)[j].$$

■

次に循環たたみ込み積と離散フーリエ変換の間の基本的な関係式を導く．

**定理 18.12** $\mathbf{a} \in \boldsymbol{K}^N$ に対して

$$\mathcal{C}_{\mathbf{a}}(\mathbf{x}) = \mathcal{F}^{-1}\left(\mathcal{F}(\mathbf{a})_{\bullet}\, \mathcal{F}(\mathbf{y})\right)$$

である．ここで $\bullet$ はアダマール積 (第 2.2.6 節参照) を表している．

この定理は次の等式から導かれる．

**定理 18.13** $\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\mathcal{F}(\mathbf{x} *_c \mathbf{y}) = \mathcal{F}(\mathbf{x})_{\bullet}\, \mathcal{F}(\mathbf{y}).$$

**証明.**

$$\mathcal{F}(\mathbf{x} *_c \mathbf{y})[k] = \sum_{n=0}^{N-1} \left( \sum_{\mu=0}^{N-1} x[\mu] y[n-\mu] \right) \omega^{-kn}$$

$$= \sum_{\mu=0}^{N-1} x[\mu] \left( \sum_{n=0}^{N-1} y[n-\mu]\, \omega^{-(n-\mu)k} \right) \omega^{-k\mu}.$$

ここで，$n - \mu = n'$ とする．このとき

$$\sum_{n=0}^{N-1} y[n-\mu]\, \omega^{-(n-\mu)k} = \sum_{n'=-\mu}^{N-1-\mu} y[n']\, \omega^{-n'k}$$

$$= \sum_{n'=-\mu}^{-1} y[n']\, \omega^{-n'k} + \sum_{n'=0}^{N-1-\mu} y[n']\, \omega^{-n'k}$$

$$= \sum_{n'=1}^{\mu} y[-n']\, \omega^{n'k} + \sum_{n'=0}^{N-1-\mu} y[n']\, \omega^{-n'k}$$

$$= \sum_{n'=1}^{\mu} y\left[N-n'\right] \omega^{-(N-n')k} + \sum_{n'=0}^{N-1-\mu} y\left[n'\right] \omega^{-n'k}$$

$$= \sum_{n=0}^{N-1} y\left[n\right] \omega^{-nk} = \mathcal{F}\left(\mathbf{y}\right)\left[k\right].$$

したがって，

$$\mathcal{F}\left(\mathbf{x} *_c \mathbf{y}\right)\left[k\right] = \sum_{\mu=0}^{N-1} x\left[\mu\right] \omega^{-k\mu} \mathcal{F}\left(\mathbf{y}\right)\left[k\right] = \mathcal{F}\left(\mathbf{x}\right)\left[k\right]\mathcal{F}\left(\mathbf{y}\right)\left[k\right]$$

となっている．■

**定理 18.12 の証明.**

$$\mathcal{C}_{\mathbf{a}}\left(\mathbf{x}\right) = \mathbf{a} *_c \mathbf{x} = \mathcal{F}^{-1}\left(\mathcal{F}\left(\mathbf{a} *_c \mathbf{x}\right)\right) = \mathcal{F}^{-1}\left(\mathcal{F}\left(\mathbf{a}\right)_{\bullet} \mathcal{F}\left(\mathbf{x}\right)\right).$$

■

**定義 18.5**　一般に $\mathbf{m} \in \boldsymbol{K}^N$ に対して，

$$T_{\mathbf{m}}(\mathbf{x}) = \mathcal{F}^{-1}\left(\mathbf{m}_{\bullet}\mathcal{F}\left(\mathbf{x}\right)\right), \quad \mathbf{x} \in \boldsymbol{K}^N$$

で定義される $\boldsymbol{K}^N$ から $\boldsymbol{K}^N$ への線形写像をシンボル $\mathbf{m}$ のフーリエ・マルチプライヤという．

**問題 18.1**

$$\mathcal{F}^{-1}\left(\mathbf{x}\right)\left[k\right] = \frac{1}{N}\mathcal{F}\left(\mathbf{x}\right)\left[-k\right] = \frac{1}{N}\mathcal{F}\left(\mathbf{x}\right)\left[N-k\right].$$

**問題 18.2**　$\mathcal{F}^{-1}\left(\mathbf{x} *_c \mathbf{y}\right) = N\,\mathcal{F}^{-1}\left(\mathbf{x}\right)_{\bullet} \mathcal{F}^{-1}\left(\mathbf{y}\right).$

**問題 18.3**　(1)　$\mathcal{F}\left(\mathbf{x}_{\bullet}\mathbf{y}\right) = \dfrac{1}{N}\mathcal{F}\left(\mathbf{x}\right) *_c \mathcal{F}\left(\mathbf{y}\right).$

(2)　$\mathcal{F}^{-1}\left(\mathbf{x}_{\bullet}\mathbf{y}\right) = \mathcal{F}^{-1}\left(\mathbf{x}\right) *_c \mathcal{F}^{-1}\left(\mathbf{y}\right).$

定理 18.12 は周波数の加工に関する必要な性質をみたすフィルタの設計の仕方を示唆している．たとえば $S$ を $\dfrac{N}{2}$ よりも小さい自然数とし，

$$\Sigma = \left\{k : \left|k - \frac{N}{2}\right| < S\right\}$$

とおき，$\mathbf{m} = \left(m\left[n\right]\right)_{n=0,\cdots,N-1} \in \boldsymbol{K}^N$ を

$$m[k]=\begin{cases}1, & k\notin\Sigma\\ 0, & k\in\Sigma\end{cases}$$

により定める．$\mathbf{a}=\mathcal{F}^{-1}(\mathbf{m})$ とすると，定理 18.12 より

$$\mathcal{C}_{\mathbf{a}}(\mathbf{x})=\mathcal{F}^{-1}(\mathbf{m}_\bullet\mathcal{F}(\mathbf{x}))$$

となっている．したがってフィルタ $\mathbf{a}$ によるフィルタリングにより，$\mathbf{x}$ の高周波部分はカットされ，低周波部分だけが活かされることになる．このようなフィルタを**低域通過フィルタ**という．特に特定の周波数領域を完全にカットし，それ以外の部分には全く手を加えないようなフィルタを**理想的フィルタ**という．

たとえば図 18.2 の一段目の左図の信号は筆者の「アー」という音声の一部をサンプリングして作ったものである (詳細は第 1.1 節参照)．この信号は 410 個のデータからなっており，それを $\mathbf{x}$ とおく．$\mathbf{x}\in\boldsymbol{K}^{410}$ である．これにあるノイズ $\boldsymbol{\varepsilon}\in\boldsymbol{K}^{410}$(白色雑音と呼ばれるもの) が入った信号 $\mathbf{x}+\boldsymbol{\varepsilon}$ のグラフが図 18.2 の一段目の右図である．

さて，このノイズは小さなすべての周波数成分からなっている．一方，図 18.2 の二段目の左図は $x$ の離散フーリエ変換の絶対値のグラフであるが，これからわかるように原信号はほとんどが低周波成分からなっている．実際，図 18.2 の二段目の右図がノイズの入った信号の離散フーリエ変換の絶対値である．そこで高周波をカットさせるフィルタを使ってフィルタリングすれば，原信号の原形をあまり損ねずにノイズを除去することができると考えられる．実際に計算機を使って計算してみよう．ここでは，$\boldsymbol{Z}_{410}=\{0,1,\cdots,409\}$, $\Sigma=\{50,51,\cdots,350\}$ とし，

$$m[k]=\begin{cases}1, & k\in\boldsymbol{Z}_{410}\setminus\Sigma\\ 0, & k\in\Sigma\end{cases}$$

により定義される $\mathbf{m}=(m[k])_{k=0,\cdots,409}\in\boldsymbol{K}^{410}$ を用いる．$\mathbf{a}=\mathcal{F}^{-1}(\mathbf{m})$ とする．このとき $\mathbf{y}=\mathcal{C}_{\mathbf{a}}(\mathbf{x}+\boldsymbol{\varepsilon})$ は図 18.2 の三段目の右図である．また三段目の左図は $\mathbf{m}_\bullet\mathcal{F}(\mathbf{x})$ の絶対値のグラフである．図 18.2 の三段目の右図を見るとノイズがある程度除去されていることがわかるであろう．

なお，ノイズ除去の方法として「**閾値**」を用いるものがある．これはノイズの入った信号 $\mathbf{x}\in\boldsymbol{K}^N$ に対して，正の数 $\gamma$ を適切にとり，もし $|\widehat{x}[k]|>\gamma$ ならば $a_k=\widehat{x}[k]$ とし，もし $|\widehat{x}[k]|\le\gamma$ ならば $a_k=0$ と定める．そして

$$\mathbf{x}'=\sum_{k=0}^{N-1}a_k\mathbf{F}_k$$

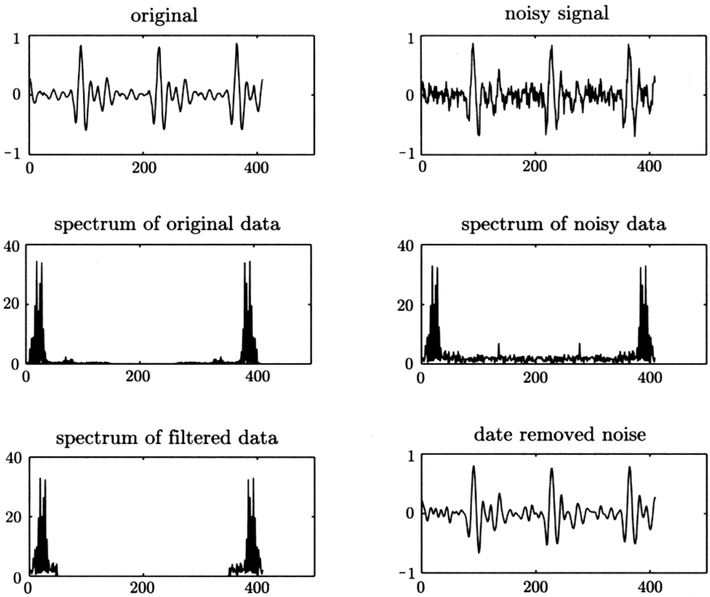



図 18.2

一段目の左図は高周波部分をカットしたもの．右図はその逆離散フーリエ変換．
二段目の左図は原信号の離散フーリエ変換の絶対値で，右図はノイズの入った信号の離散フーリエ変換の絶対値．
三段目の左図は高周波部分をカットしたもの．右図はその逆離散フーリエ変換．

と定義する．$\mathbf{x}'$ はしばしばノイズの除去された信号となっている．$\gamma$ を閾値という．

本書では理想的なフィルタの例をあげたが，このほかにもいろいろなタイプのフィルタが知られている．フィルタの詳細についてはディジタル信号処理の本などを参照してほしい (たとえば [OS] を参照).

## 18.4　循環相関積

循環たたみ込み積と似ている演算に次の**循環相関積**がある．ウェーブレットではたたみ込み積と循環相関積の両方を用いて解析が進められる．本節では循環相関積について述べる．$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$x \bigstar y[k] = \sum_{n=0}^{N-1} x[n]y[n+k], \quad k \in \boldsymbol{Z}_N$$

とし，

$$\mathbf{x}\bigstar\mathbf{y} = \begin{pmatrix} x\bigstar y[0] \\ \vdots \\ x\bigstar y[N-1] \end{pmatrix}$$

とする．これを $\mathbf{x}$ と $\mathbf{y}$ の**循環相関積**という．循環たたみ込み積は可換であるが，循環相関積は可換ではない．実際，$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して

$$x\bigstar y[k] = y\bigstar x[-k]$$

である．なぜならば，$n + k = m$ とおくと，

$$x\bigstar y[k] = \sum_{n=0}^{N-1} x[n]y[n+k] = \sum_{m=k}^{N-1+k} y[m]x[m-k] = y\bigstar x[-k]$$

となっているからである．

いま $\mathbf{x} = (x\,[n])_{n\in\boldsymbol{Z}_N} \in \boldsymbol{K}^N$ に対して

$$x^{\vee}\,[n] = x\,[-n] \tag{18.7}$$

とし，

$$\mathbf{x}^{\vee} = (x^{\vee}\,[n])_{n\in\boldsymbol{Z}_N}$$

と定義する．次の関係式が成り立っている．

**定理 18. 14**　$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\mathbf{x}\bigstar\mathbf{y} = (\mathbf{y}\bigstar\mathbf{x})^{\vee},$$
$$\mathbf{x}\bigstar\mathbf{y} = \mathbf{x}^{\vee} *_c \mathbf{y}.$$

**証明.** 前半の等式はすでに証明したので，後半を示す．$m = k + n$ と変数変換して

$$\begin{aligned} x\bigstar y\,[k] &= \sum_{n=0}^{N-1} x\,[n]\,y\,[k+n] = \sum_{m=k}^{k+N-1} x\,[m-k]\,y\,[m] \\ &= \sum_{m=0}^{N-1} x\,[m-k]\,y\,[m] = \sum_{m=0}^{N-1} x^{\vee}\,[k-m]\,y\,[m] = y *_c x^{\vee}\,[k] \\ &= x^{\vee} *_c y\,[k]. \end{aligned}$$

■

$\mathbf{a} \in \boldsymbol{K}^N$ とする．このとき循環相関積により線形写像

$$S_{\mathbf{a}}(\mathbf{x}) = \mathbf{a} \bigstar \mathbf{x}$$

が定義される．定理 18.14 より，この線形写像はたたみ込み作用素と

$$S_{\mathbf{a}}(\mathbf{x}) = \mathcal{C}_{\mathbf{a}^{\vee}}(\mathbf{x})$$

の関係にある．$\mathbf{a} = (a_n)_{n=0,\cdots,N-1} \in \boldsymbol{K}^N$ に対して $S_{\mathbf{a}}$ の表現行列が巡回行列

$$\operatorname{circ}(a_0, \cdots, a_{N-1}) = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{N-1} \\ a_{N-1} & a_0 & a_1 & \cdots & a_{N-2} \\ a_{N-2} & a_{N-1} & a_0 & \cdots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$$

であることが次に示すように容易にわかる．

**定理 18. 15**

$$S_{\mathbf{a}}(\mathbf{x}) = \operatorname{circ}(a_0, a_1 \cdots, a_{N-1})\,\mathbf{x}. \tag{18.8}$$

したがって，フィルタ $\mathbf{a}$ によるフィルタリングの表現行列による表示は

$$\mathcal{C}_{\mathbf{a}}(\mathbf{x}) = \operatorname{circ}(a_0, a_{N-1}, \cdots, a_1)\,\mathbf{x} \tag{18.9}$$

である．

**証明.**

$$(18.8) \text{ の右辺} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{N-1} \\ a_{-1} & a_0 & a_1 & \cdots & a_{N-2} \\ a_{-2} & a_{-1} & a_0 & \cdots & a_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{-N+1} & a_{-N+2} & a_{-N+3} & \cdots & a_0 \end{pmatrix} \begin{pmatrix} x[0] \\ x[1] \\ x[2] \\ \vdots \\ x[N-1] \end{pmatrix}$$

$$
= \begin{pmatrix} \sum_{k=0}^{N-1} a_k x[k] \\ \sum_{k=0}^{N-1} a_{k-1} x[k] \\ \sum_{k=0}^{N-1} a_{k-2} x[k] \\ \vdots \\ \sum_{k=0}^{N-1} a_{k-(N-1)} x[k] \end{pmatrix} = \begin{pmatrix} \sum_{k=0}^{N-1} a_k x[k] \\ \sum_{k=0}^{N-1} a_k x[k+1] \\ \sum_{k=0}^{N-1} a_k x[k+2] \\ \vdots \\ \sum_{k=0}^{N-1} a_k x[k+N-1] \end{pmatrix}
$$

$$
= \mathbf{a} \bigstar \mathbf{x}.
$$

■

## 18.5 フーリエ行列と巡回行列

本節では，巡回行列がフーリエ行列により対角化できることを証明する．なお，すでに示したように，フーリエ行列はユニタリー行列であることを思い出してほしい．

まず巡回行列はスカラー倍と加法について，

$$
\alpha \operatorname{circ}(a_0, \cdots, a_{N-1}) = \operatorname{circ}(\alpha a_0, \cdots, \alpha a_{N-1}), \quad (18.10)
$$

$$
\operatorname{circ}(a_0, \cdots, a_{N-1}) + \operatorname{circ}(b_0, \cdots, b_{N-1}) = \operatorname{circ}(a_0 + b_0, \cdots, a_{N-1} + b_{N-1}) \quad (18.11)
$$

をみたしていることが容易に証明できることに注意しておく．巡回行列の基本的な性質として，すべての巡回行列は

$$
\Pi = \operatorname{circ}(0, 1, 0, \cdots, 0)
$$

という特別な巡回行列を用いて記述することができることを証明する．

**定理 18.16**

$$
\begin{aligned}
&\operatorname{circ}(a[0], \cdots, a[N-1]) \\
&\quad = a[0]\ I_N + a[1]\ \Pi + \cdots + a[N-1]\ \Pi^{N-1}. \qquad (18.12)
\end{aligned}
$$

**証明．** (18.10), (18.11) より

$$
\begin{aligned}
&\mathrm{circ}\,(a\,[0]\,,\cdots,a\,[N-1]) \\
&\quad= \mathrm{circ}\,(a\,[0]\,,0,\cdots,0) + \mathrm{circ}\,(0,a\,[1]\,,0,\cdots,0) \\
&\qquad+ \cdots + \mathrm{circ}\,(0,\cdots 0,a\,[N-1]) \\
&\quad= a\,[0]\;I_N + a\,[1]\;\mathrm{circ}\,(0,1,0,\cdots,0) \\
&\qquad+ a\,[2]\;\mathrm{circ}\,(0,0,1,0,\cdots,0) + \cdots + a\,[N-1]\;\mathrm{circ}\,(0,\cdots 0,1)\,.
\end{aligned}
$$

一方

$$
\begin{aligned}
&\Pi^2 = \mathrm{circ}\,(0,0,1,\cdots,0)\,, \quad \Pi^3 = (0,0,0,1,0,\cdots,0)\,, \quad \cdots, \\
&\Pi^{N-1} = (0,\cdots,0,1)\,, \qquad \Pi^N = I_N
\end{aligned}
$$

となっていることが容易に確認できる．したがって，(18.12) が得られる．■

巡回行列の対角化を得るために，まず $\Pi$ がフーリエ行列により対角化できることを示す．

$$
\omega = \exp\left(\frac{2\pi i}{N}\right)
$$

とし，

$$
\Omega = \mathrm{diag}\left(1,\omega,\omega^2,\cdots,\omega^{N-1}\right)
$$

とする．このとき次のことが成り立つ．

**補題 18. 17**　$F$ を $N$ 次フーリエ行列とすると

$$
\Pi = F^*\Omega F
$$

である．

**証明．** $F^* = (g_{mn})_{\substack{n=0,\cdots,N-1\\ m=0,\cdots,N-1}}$ とすると $g_{nm} = \dfrac{1}{\sqrt{N}}\omega^{nm}$ である．$\Omega = (\omega^n\delta_{mn})$ であるから

$$
F^*\Omega = \left(\sum_{\mu=0}^{N-1} g_{m\mu}\omega^{\mu}\delta_{\mu n}\right) = (g_{mn}\omega^n) = \frac{1}{\sqrt{N}}\left(\omega^{(m+1)n}\right).
$$

したがって，

$$
F^*\Omega F = \frac{1}{N}\left(\sum_{\mu=0}^{N-1}\omega^{(m+1)\mu}\omega^{-\mu n}\right) = \frac{1}{N}\left(\sum_{\mu=0}^{N-1}\omega^{(m-n+1)\mu}\right).
$$

ここで，

$$\frac{1}{N}\sum_{\mu=0}^{N-1}\omega^{(m-n+1)\mu}=\begin{cases}1, & m-n+1=0\\ 0, & m-n+1\neq 0\end{cases}$$

であるから，$F^*\Omega F=\Pi$ が得られる．■

この補題および $FF^*=F^*F=I_N$ より $\Pi^n=F^*\Omega^n F$ であるから，

$$\begin{aligned}
&\operatorname{circ}(a[0],\cdots,a[N-1])\\
&\quad=a[0]\,I_N+a[1]\,\Pi+\cdots+a[N-1]\,\Pi^{N-1}\\
&\quad=a[0]\,F^*I_NF+a[1]\,F^*\Omega F+\cdots+a[N-1]\,F^*\Omega^{N-1}F\\
&\quad=F^*\left(a[0]\,I_N+a[1]\,\Omega+\cdots+a[N-1]\,\Omega^{N-1}\right)F
\end{aligned}$$

となっている．ここで

$$Z_{\mathbf{a}}(z)=a[0]+a[1]\,z+\cdots+a[N-1]\,z^{N-1},\quad z\in\boldsymbol{C} \tag{18.13}$$

とおくと，

$$\begin{aligned}
&a[0]\,I_N+a[1]\,\Omega+\cdots+a[N-1]\,\Omega^{N-1}\\
&\quad=\operatorname{diag}\left(Z_{\mathbf{a}}(1),Z_{\mathbf{a}}(\omega),\cdots,Z_{\mathbf{a}}\left(\omega^{N-1}\right)\right)
\end{aligned}$$

と表せる．なお，$Z_{\mathbf{a}}(z^{-1})$ はベクトル $\mathbf{a}=(a[j])\in\boldsymbol{K}^N$ に対する $z$ **変換**と呼ばれるものである．以上のことをまとめると，巡回行列の対角化について次の定理が得られる．

**定理 18.18**　$\mathbf{a}=(a[j])\in\boldsymbol{K}^N$ とし，$F$ を $N$ 次フーリエ行列とする．このとき

$$F\operatorname{circ}(a[0],\cdots,a[N-1])\;F^*=\operatorname{diag}\left(Z_{\mathbf{a}}(1),Z_{\mathbf{a}}(\omega),\cdots,Z_{\mathbf{a}}\left(\omega^{N-1}\right)\right)$$

である．

定理 18.18 を用いれば次のことが容易に証明できる．

**系 18.19**　$\operatorname{circ}(a[0],\cdots,a[N-1])$ の固有値は $Z_{\mathbf{a}}(1)$, $Z_{\mathbf{a}}(\omega)$, $\cdots$, $Z_{\mathbf{a}}\left(\omega^{N-1}\right.$ である．

**系 18.20**　$\det\operatorname{circ}(a[0],\cdots,a[N-1])=Z_{\mathbf{a}}(1)\,Z_{\mathbf{a}}(\omega)\cdots Z_{\mathbf{a}}\left(\omega^{N-1}\right)$.

最後に，任意の対角行列 $\Delta$ に対して

$$F\operatorname{circ}(a[0],\cdots,a[N-1])\;F^*=\Delta$$

をみたす巡回行列が存在することを示す．$\mathbf{m}=(m[n])\in\boldsymbol{K}^N$ に対して，

$$\Delta_{\mathbf{m}} = \mathrm{diag}(m[0], \cdots, m[N-1])$$

とする.

**命題 18.21**　任意の $\mathbf{m} = (m[j]) \in \boldsymbol{K}^N$ に対して，$\mathbf{a} = \dfrac{1}{N}\mathcal{F}(\mathbf{m})$ とすると

$$\begin{aligned}\Delta_{\mathbf{m}} &= \mathrm{diag}\left(Z_{\mathbf{a}}(1), Z_{\mathbf{a}}(\omega), \cdots, Z_{\mathbf{a}}\left(\omega^{N-1}\right)\right) \\ &= F\,\mathrm{circ}\,(a[0], \cdots, a[N-1])\; F^*\end{aligned}$$

が成り立つ.

**証明.** $Z_{\mathbf{a}}(1) = m[0]$, $Z_{\mathbf{a}}(\omega) = m[1]$, $\cdots$, $Z_{\mathbf{a}}(\omega^{N-1}) = m[N-1]$ をみたすような $\mathbf{a}$ を求めればよい．これが成り立つことは，(18.13) の定義から

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} a[0] \\ a[1] \\ \vdots \\ a[N-1] \end{pmatrix} = \begin{pmatrix} m[0] \\ m[1] \\ \vdots \\ m[N-1] \end{pmatrix}$$

が成り立つことと同値である．すなわち $\sqrt{N}F^*\mathbf{a} = \mathbf{m}$ が成り立つことと同値である．$F$ はユニタリー行列であるから，求める $\mathbf{a}$ は，$\mathbf{a} = \dfrac{1}{\sqrt{N}}F\mathbf{m} = \dfrac{1}{N}\mathcal{F}(m)$ で与えられる．■

この定理から次のことがわかる．任意に $\mathbf{m} = (m[j]) \in \boldsymbol{K}^N$ をとり，信号 $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$\mathcal{F}^{-1}(\mathbf{m}_{\bullet}\mathcal{F}(\mathbf{x}))$$

により $\mathbf{x}$ の周波数成分を整形する．この操作は，$\mathbf{a} = \dfrac{1}{N}\mathcal{F}(\mathbf{m})$ とすると次のように表される.

$$\begin{aligned}\mathcal{F}^{-1}(\mathbf{m}_{\bullet}\mathcal{F}(\mathbf{x})) &= \frac{1}{\sqrt{N}}F^*\Delta_{\mathbf{m}}\sqrt{N}F\mathbf{x} \\ &= F^*\mathrm{diag}\left(Z_{\mathbf{a}}(1), Z_{\mathbf{a}}(\omega), \cdots, Z_{\mathbf{a}}\left(\omega^{N-1}\right)\right)F \\ &= \mathrm{circ}\,(a[0], \cdots, a[N-1])\,\mathbf{x} \\ &= S_{\mathbf{a}}(\mathbf{x}) \\ &= \mathcal{C}_{\mathbf{a}^\vee}(\mathbf{x}).\end{aligned}$$

**問題 18.4**　$A, B$ が $N$ 次の巡回行列ならば，$AB$ も $N$ 次の巡回行列であるこ

とを示せ.

## 18.6 スペクトログラム

いま楽音が 44100 [Hz] のサンプリング周波数で記録されているとする (たとえば音楽 CD など). もし演奏時間 1 時間の楽曲が録音されている場合, 楽音に関連する実質的なデータ数は $44100 \times 60 \times 60 = 158760000$ 個である. したがって, この楽曲を 158760000 次元のベクトルと考えて, 離散フーリエ変換をすることは理論上可能である. これによりこのベクトルがどのような周波数成分をもつかがわかる. しかしこれではどの時間にどのくらいの周波数の音が発せられたかがわからない. そこで離散フーリエ変換の代わりにスペクトログラムという方法を用いることが多い. これは各時刻の近傍のみで局所的に離散フーリエ変換していくというもので, これを使うとどの時間にどのくらいの周波数の音が発せられたかが“ある程度”わかる.

いまアナログ信号 $x_c(t)$ が $f_s$ [Hz] のサンプリング周波数でサンプリングされ, 離散時間信号 $x[n]\ (n = 0, \cdots, N-1)$ が得られたとする. この信号を局所的に扱うために, ウィンドウと呼ばれるベクトル $h[n]$ を用意しておく. 本節ではウィンドウは

$$\{n : h[n] \neq 0\} \subset [0, L-1], \quad L < N$$

をみたしているとする. $L$ をウィンドウの長さという. よく使われるウィンドウの例としては次のようなものがある.

**例 18.1**

$$h[n] = 0.5\left(1 - \cos\left(2\pi \frac{n}{L-1}\right)\right), \quad n = 0, \cdots, L-1$$

を $L$ 点ハニング・ウィンドウという.

**例 18.2** $\alpha \geq 2,\ n = 0, \cdots, L-1$ に対して,

$$h[n] = \exp\left(-\frac{1}{2}\left(\alpha \frac{n - \dfrac{L-1}{2}}{\dfrac{L}{2}}\right)^2\right)$$

とする. これをガウシアン・ウィンドウという.

さて, $h = (h[n]) \in \boldsymbol{K}^L$ をウィンドウとする. ただし $2 \leq L < N$ であるとする.

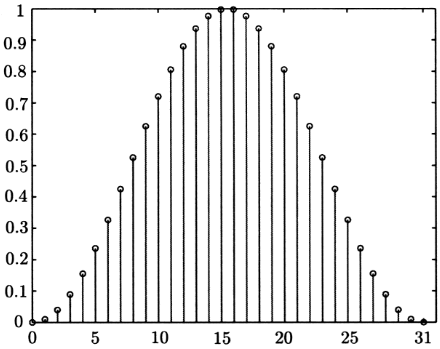



**図 18.3**　32 点ハニング・ウィンドウ

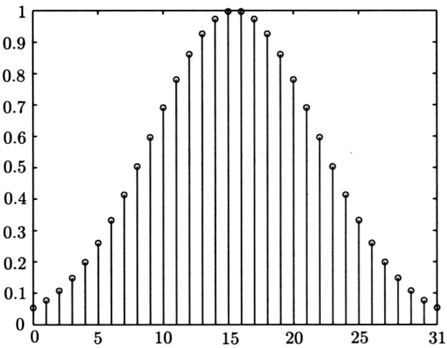



**図 18.4**　ガウシアン・ウィンドウ，$\alpha = 2,\, L = 32.$

$N_0$ を $N_0 < L$ をみたす自然数とする．スペクトログラムはウィンドウを時間進行方向に移動させて作成するのだが，その際，隣接するウィンドウが重なるように移動させていく．$N_0$ をウィンドウの重なる点の数を表すものとする．$s = L - N_0$ とし，

$$n_0 = 1 + \left\lfloor \frac{N-L}{s} \right\rfloor$$

とする．ここで実数 $a$ に対して $\lfloor a \rfloor$ は $a$ を超えない最大の整数 (下方整数部分) を表す．$j = 0, \cdots, n_0 - 1$ とし，$m_j = js$ とする．

$$m_0 = 0,\ m_1 = s,\ m_2 = 2s, \cdots,\ m_{n_0-1} = (n_0 - 1)\, s$$

である．また $\mathbf{x}_j \in \boldsymbol{K}^L$ を $\mathbf{x}_j^T = (x[m_j], \cdots x[m_j + L - 1])$ と定める．このとき，$\mathbf{x}_j$ の定義域は

$$
\begin{aligned}
\mathbf{x}_0 \text{ の定義域 } &: 0, \cdots, L-1 \\
\mathbf{x}_1 \text{ の定義域 } &: s, \cdots, s+L-1 \\
\mathbf{x}_2 \text{ の定義域 } &: 2s, \cdots, 2s+L-1 \\
&\vdots \\
\mathbf{x}_{n_0-1} \text{ の定義域 } &: (n_0-1)s, \cdots, (n_0-1)s+L-1
\end{aligned}
$$

である．ここで

$$(n_0-1)s+L-1 \leq \left(\frac{N-L}{s}\right)s+L-1 = N-1$$

であることに注意しておく．いま $L \leq M \leq N$ なる偶数 $M$ をとり固定する．$\mathbf{h}_\bullet \mathbf{x}_j \in \boldsymbol{K}^L$ とみなせるが，これに対して $(\mathbf{h}_\bullet \mathbf{x}_j)_M \in \boldsymbol{K}^M$ を次のように定義する．

$$
\begin{aligned}
&(\mathbf{h}_\bullet \mathbf{x}_j)_L^T \\
&\quad = (h[0]x[m_j],\, h[1]\, x[m_j+1], \cdots, h[L-1]x[m_j+L-1], 0, \cdots, 0).
\end{aligned}
$$

そして，$\mathbf{X}_j = \mathcal{F}_M\left((\mathbf{h}_\bullet \mathbf{x}_j)_M\right)$ とおく．$\mathbf{X}_j = (X(k,j))_{k=0,\cdots,M-1} \in \boldsymbol{K}^M$ とおき，$\left(\dfrac{M}{2}+1, n_0\right)$ 行列

$$
\begin{pmatrix}
\left|X\left(\frac{M}{2},0\right)\right| & \left|X\left(\frac{M}{2},1\right)\right| & \cdots & \left|X\left(\frac{M}{2},n_0-1\right)\right| \\
\vdots & \vdots & & \vdots \\
|X(1,0)| & |X(1,1)| & & |X(1,n_0-1)| \\
|X(0,0)| & |X(0,1)| & \cdots & |X(0,n_0-1)|
\end{pmatrix}
$$

を定義する．これを $\mathbf{x}$ の**スペクトログラム**という．スペクトログラムはウィンドウ $h$ と重複数 $N_0$ および $M$ のとり方により値が変わってくる．$\mathbf{X}_j$ はおおよそ

$$\left(m_j+\frac{L}{2}\right)\frac{1}{f_s} = \left(js+\frac{L}{2}\right)\frac{1}{f_s} \quad (\text{秒})$$

の近傍における $k\dfrac{f_s}{M}$ [Hz] での周波数成分の絶対値 $|X(k,j)|$, $k=0,\cdots,\dfrac{M}{2}$ を記述している．$|X(k,j)|$ の値を 0 が白，そして値が大きくなるにつれて灰色から黒へと色の濃淡で描くことにより，スペクトログラムをグラフにすることができる．横軸は時間軸，縦軸は周波数を表す．

なお場合によっては $|X(k,j)|$ の代わりに $\log|X(k,j)|$ を用いてスペクトログラムを定義することもある．

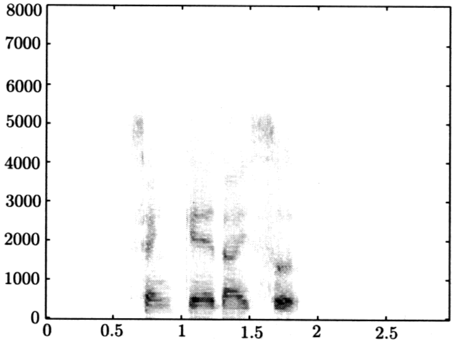



**図 18.5**

スペクトログラムの例．筆者が「線形代数」と言った音声．サンプリング周波数は 16 [kHz]．長さ 256 のハニング・ウィンドウで，$M = 256, N_0 = 250$.

# 第19章　離散ウェーブレットへの応用

本章ではウェーブレット解析を線形代数の立場から論ずる．標語的にいえば，ウェーブレット解析とはスケーリング・フィルタと呼ばれる低域通過フィルタとウェーブレット・フィルタと呼ばれる帯域通過フィルタを用いたサブバンド信号処理とみなすことができる．まずこれらの用語を定義していきたい．

本章では $\boldsymbol{K}^N$ の元 $\mathbf{x} = (x\,[n])_{n=0,\cdots,N-1}$ を長さ $N$ の**信号**という．信号処理では一般に，与えられた信号 $\mathbf{x} = (x\,[n])_{n=0,\cdots,N-1}$ を，便宜上 $n < 0$ に対する値 $x[n]$ と $n > N-1$ に対する値 $x[n]$ を適当な規則で定めることにより，信号の定義域を拡張して処理することが多い．信号の定義域の拡張の仕方にはいくつかの方法があるが，ここでは第 18.3 節で述べた線形代数の枠組みで捉えやすい規約 18.9 に従った周期拡張を考えることにする．すなわち与えられた信号 $\mathbf{x} = (x\,[n])_{n=0,\cdots,N-1} \in \boldsymbol{K}^N$ は $j = 0,\cdots,N-1$, $k = \pm1, \pm2, \cdots$ に対して，

$$x\,[j + kN] = x\,[j]$$

と周期的に拡張されているものとする．

なお，この規約の元に

$$\tau_k \mathbf{x} = \begin{pmatrix} x\,[-k] \\ x\,[1-k] \\ \vdots \\ x\,[N-1-k] \end{pmatrix} \in \boldsymbol{K}^N$$

と定義し，これを $k$-**シフト変換**という．

**以下，本章では $N$ を正の偶数とし，$N_1 = \dfrac{N}{2}$ とする．** $N$ としては大きな偶数を想定している．

**ポイント 19**　ウェーブレットは 1980 年代に考えられた新しい数学的な方法で

ある．ウェーブレットを線形代数の枠組みの中で論ずることは，有限個の数からなるデータに対するウェーブレット解析をする際の基礎になる．なお本書でのウェーブレットの扱いは，従来の理論に比べて，数学的に一般化してある部分もある．

## 19.1　準備

一般にウェーブレットは二種類のベクトルからなる．一つはスケーリング・フィルタと呼ばれるもので，もう一つはウェーブレット・フィルタと呼ばれるものである．一般的な定義を行う前に，まず具体例を見ることにしよう．ウェーブレットの最も単純な例はハール・ウェーブレットと呼ばれるものである．

$$\mathbf{u}_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\\vdots\\0\end{pmatrix}, \quad \mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\\vdots\\0\end{pmatrix} \in \boldsymbol{K}^N$$

とし，$\mathbf{u}_0$ をハール・スケーリング・フィルタ，$\mathbf{u}_1$ をハール・ウェーブレット・フィルタという．どちらも $n=0,1$ でのみ変動があり，$n \geq 2$ では定常的に 0 となっている (図 19.1 参照).

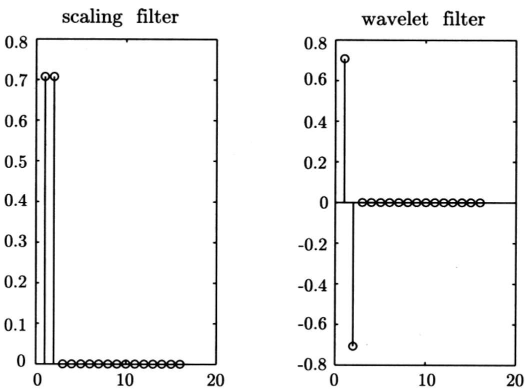



**図 19.1**

左図はハール・スケーリング・フィルタ，右図はハール・ウェーブレット・フィルタ．

ウェーブレットはこのほかにもたくさんの例が知られているが，たとえばドブ

シーにより発見された図 19.2 のようなウェーブレットもある (詳細は第 19.5 節で解説する). 本書ではこれをドブシー 2 と呼ぶことにする.

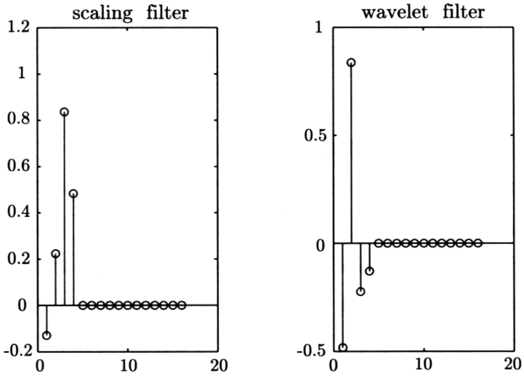



**図 19.2**

左図はドブシー 2 スケーリング・フィルタ，右図はドブシー 2 ウェーブレット・フィルタ.

さてこれらのフィルタの周波数特性を見るために離散フーリエ変換を施してみよう. $N = 16$ とする. 図 19.3 はハール・スケーリング・フィルタとハール・ウェーブレット・フィルタを離散フーリエ変換し，さらに絶対値をとったものである. また図 19.4 はドブシー 2 の離散フーリエ変換の絶対値のグラフである. どちらにも特徴的なことは，スケーリング・フィルタは高周波の部分 (すなわち $n = 8$ 付近) で減少し，低周波の部分 (すなわち $n = 0$ ならびに $n = 15$ の付近) では増加していることである. 一方，ウェーブレット・フィルタでは逆になっている. このことはスケーリング・フィルタは低域通過フィルタ，ウェーブレット・フィルタは高域通過フィルタとなっていることを意味している.

スケーリング・フィルタ，ウェーブレット・フィルタの一般的な定義は次のものである.

**定義 19.1** $\mathbf{u}_0 = (u_0\,[n])_{n=0}^{N-1} \in \boldsymbol{K}^N$ が**スケーリング・フィルタ**であるとは,

$$\mathcal{F}_N\,(\mathbf{u}_0)\,[0] = \sqrt{2}, \quad \mathcal{F}_N\,(\mathbf{u}_0)\,[N_1] = 0 \tag{19.1}$$

が成り立つことである. また, $\mathbf{u}_1 = (u_1\,[n])_{n=0}^{N-1} \in \boldsymbol{K}^N$ が**ウェーブレット・フィルタ**であるとは,

$$\mathcal{F}_N\,(\mathbf{u}_1)\,[0] = 0, \quad \mathcal{F}_N\,(\mathbf{u}_1)\,[N_1] = \sqrt{2} \tag{19.2}$$

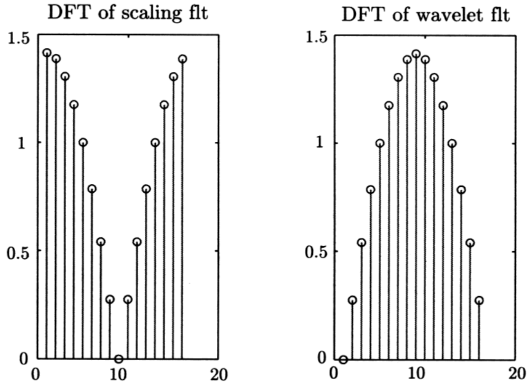



図 19.3

が成り立つことである．

離散フーリエ変換の定義に基づいて計算すれば，スケーリング・フィルタであるための必要十分条件は

$$\sum_{n=0}^{N-1} u_0[n] = \sqrt{2}, \quad \sum_{n=0}^{N-1} (-1)^n u_0[n] = 0, \tag{19.3}$$

また，ウェーブレット・フィルタであるための必要十分条件は

$$\sum_{n=0}^{N-1} u_1[n] = 0, \quad \sum_{n=0}^{N-1} (-1)^n u_1[n] = \sqrt{2} \tag{19.4}$$

と表わすことができる．

ウェーブレット解析ではこれを用いて，与えられた信号に対して次のような処理を行う．簡単のためスケーリング・フィルタもウェーブレット・フィルタも実数値であるとする．まず与えられた信号 $\mathbf{x} \in \boldsymbol{K}^N$ をスケーリング・フィルタ $\mathbf{u}_0$ とウェーブレット・フィルタ $\mathbf{u}_1$ によりフィルタリングする．ただし便宜上，循環たたみ込み積ではなく循環相関積を使う．

$$\mathbf{x}_L = \mathbf{u}_0 \bigstar \mathbf{x},$$
$$\mathbf{x}_H = \mathbf{u}_1 \bigstar \mathbf{x}$$

とおく．この二つの信号は $\mathbf{x}_L$ が $\mathbf{x}$ の低周波部分を抜き出したもの，また $\mathbf{x}_H$ は $\mathbf{x}$ の帯域部分 (または高周波部分) を抜き出したものである．ここで $\mathbf{x}$ のデータ数が $N$ 個であるのに対して，$\mathbf{x}_L$ と $\mathbf{x}_H$ を合わせるとデータ数は $2N$ 個になっていることに注

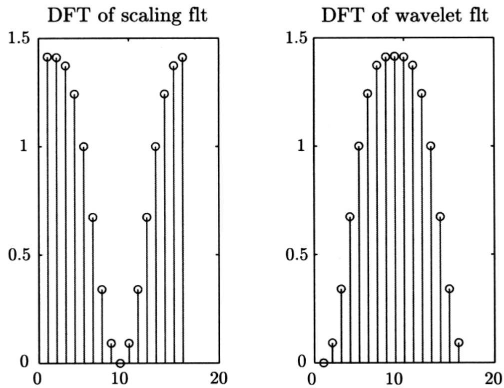



図 19.4

意する．そこでデータ数を増加させないためにサブサンプリングと呼ばれるデータの縮約を行う．サブサンプリングとはデータの一部分を一定の規則に従って抜き出す処理であるが，たとえば次のようなものを考える．$\mathbf{y} = (y[n])_{n=0,\cdots,N-1} \in \boldsymbol{K}^N$ に対して，

$$(\downarrow 2)\,\mathbf{y} = (y[2n])_{n=0,\cdots,\frac{N}{2}-1}$$

とする．$(\downarrow 2)\,\mathbf{y} \in \boldsymbol{K}^{\frac{N}{2}}$ である．そこで $\mathbf{x}_L$ と $\mathbf{x}_H$ の代わりに $(\downarrow 2)\,\mathbf{x}_L$ と $(\downarrow 2)\,\mathbf{x}_H$ に対して信号処理を行うことにすれば，データ数は変化しない．$(\downarrow 2)\,\mathbf{x}_L$ と $(\downarrow 2)\,\mathbf{x}_H$ をサブバンド信号という．

さて，$(\downarrow 2)\,\mathbf{x}_L$ と $(\downarrow 2)\,\mathbf{x}_H$ に必要な処理を加えて，次にこれらの信号を適切な方法で合成してもとの長さの信号に戻す．たとえばデータ圧縮の復号，ノイズ除去，高周波強調などではその必要がある．その際，合成の方法をどのようにすればよいかというと，もしサブバンド信号に何も処理を加えずに合成した場合，原信号 $\mathbf{x}$ と合成したものが一致していることが望ましい．そのような合成方法は存在するのだろうか？　そしてそれは具体的にどのような手順で行われるのであろうか？　この問題については次節で述べることにする．なお，サブサンプリングしないままデータを加工する方法もある．

本論に入る前に，後の考察で必要になるいくつか公式を証明する．次の定理は循環相関積，循環たたみ込み積を行列を用いて表すものである．

**定理 19.1**　$\mathbf{x} = (x[n])_{n\in \boldsymbol{Z}_N}$ に対して，

$$X = (\tau_0 \mathbf{x},\ \tau_1 \mathbf{x},\ \cdots, \tau_{N-1} \mathbf{x}) \in \boldsymbol{K}^{N \times N}$$

とする．このとき

$$X^T \mathbf{y} = \mathbf{x} \bigstar \mathbf{y},$$
$$X \mathbf{y} = \mathbf{x} *_c \mathbf{y}.$$

**証明.** まず

$$\sum_{n=0}^{N-1} \tau_j x[n]\, y[n] = \sum_{n=0}^{N-1} x[n-j]\, y[n] = \sum_{n=0}^{N-1} x[n]\, y[n+j] = x \bigstar y[j],$$

$$\sum_{n=0}^{N-1} \tau_n x[j]\, y[n] = \sum_{n=0}^{N-1} x[j-n]\, y[n] = x *_c y[j]$$

に注意しておく．定義より

$$X^T \mathbf{y} = \begin{pmatrix} \tau_0 x[0] & \tau_0 x[1] & \cdots & \tau_0 x[N-1] \\ \tau_1 x[0] & \tau_1 x[1] & \cdots & \tau_1 x[N-1] \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{N-1} x[0] & \tau_{N-1} x[1] & \cdots & \tau_{N-1} x[N-1] \end{pmatrix} \begin{pmatrix} y[0] \\ y[1] \\ \vdots \\ y[N-1] \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{n=0}^{N-1} \tau_0 x[n]\, y[n] \\ \displaystyle\sum_{n=0}^{N-1} \tau_1 x[n]\, y[n] \\ \vdots \\ \displaystyle\sum_{n=0}^{N-1} \tau_{N-1} x[n]\, y[n] \end{pmatrix} = \begin{pmatrix} x \bigstar y[0] \\ x \bigstar y[1] \\ \vdots \\ x \bigstar y[N-1] \end{pmatrix} = \mathbf{x} \bigstar \mathbf{y}.$$

また

$$
X\mathbf{y} = \begin{pmatrix} \tau_0 x[0] & \tau_1 x[0] & \cdots & \tau_{N-1} x[0] \\ \tau_0 x[1] & \tau_1 x[1] & \cdots & \tau_{N-1} x[1] \\ \vdots & \vdots & \ddots & \vdots \\ \tau_0 x[N-1] & \tau_1 x[N-1] & \cdots & \tau_{N-1} x[N-1] \end{pmatrix} \begin{pmatrix} y[0] \\ y[1] \\ \vdots \\ y[N-1] \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{n=0}^{N-1} \tau_n x[0]\, y[n] \\ \sum_{n=0}^{N-1} \tau_n x[1]\, y[n] \\ \vdots \\ \sum_{n=0}^{N-1} \tau_n x[N-1]\, y[n] \end{pmatrix} = \begin{pmatrix} x *_c y[0] \\ x *_c y[1] \\ \vdots \\ x *_c y[N-1] \end{pmatrix} = \mathbf{x} *_c \mathbf{y}.
$$

■

**問題 19. 1**　次を証明せよ．$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$
\mathbf{x} \bigstar \mathbf{y} = \tau_1 (\mathbf{x} \bigstar \tau_{-1} \mathbf{y}),
$$
$$
\mathbf{x} *_c \mathbf{y} = \tau_1 (\mathbf{x} *_c \tau_1 \mathbf{y}).
$$

離散フーリエ変換とたたみ込み積，相関との間の関係式を述べておく．

**定理 19. 2**　(1)　$\mathcal{F}_N(\mathbf{x} *_c \mathbf{y}) = \mathcal{F}_N(\mathbf{x})_\bullet \mathcal{F}_N(\mathbf{y})$.　($\bullet$ はアダマール積を表す)

(2)　$\mathcal{F}_N(\mathbf{x}^\vee) = \overline{\mathcal{F}_N(\overline{\mathbf{x}})}$.　($\mathbf{x}^\vee$ の定義は (18.7)，$-$ の定義は (2.5) 参照)

(3)　$\mathcal{F}_N(\mathbf{x} \bigstar \mathbf{y}) = \overline{\mathcal{F}_N(\overline{\mathbf{x}})}_\bullet \mathcal{F}_N(\mathbf{y})$.

**証明．** (1) はすでに証明した．

(2)

$$
\mathcal{F}_N(\mathbf{x}^\vee)[n] = \sum_{k=0}^{N-1} x[-k] \exp\left(-2\pi i \frac{kn}{N}\right) = \sum_{k=0}^{N-1} x[k] \exp\left(2\pi i \frac{kn}{N}\right)
$$
$$
= \overline{\sum_{k=0}^{N-1} \overline{x[k]} \exp\left(-2\pi i \frac{kn}{N}\right)} = \overline{\mathcal{F}_N(\overline{\mathbf{x}})[n]}.
$$

(3)

$$
\mathcal{F}_N(\mathbf{x} \bigstar \mathbf{y}) = \mathcal{F}_N(\mathbf{x}^\vee *_c \mathbf{y}) = \mathcal{F}_N(\mathbf{x}^\vee)_\bullet \mathcal{F}_N(\mathbf{y}) = \overline{\mathcal{F}_N(\overline{\mathbf{x}})}_\bullet \mathcal{F}_N(\mathbf{y}).
$$

■

サブサンプリングの行列による表示をしておく．$L$ を $L \leq N$ なる自然数とする．$L$ 個の整数を

$$0 \leq k_0 < k_1 < \cdots < k_{L-1} \leq N-1$$

をみたすようにとり，

$$J = \{k_0,\, k_1,\, \cdots,\, k_{L-1}\} \tag{19.5}$$

とする．

**定義 19.2**　$\mathbf{x} = (x[n])_{n \in \boldsymbol{Z}_N} \in \boldsymbol{K}^N$ に対して

$$(\downarrow J)\,\mathbf{x} = (x[k_j])_{j \in \boldsymbol{Z}_L} \in \boldsymbol{K}^L$$

と定義する．$(\downarrow J)$ を $J$ に関する**サブサンプラー**という．

特に，$K$ を $K < N$ なる自然数で，$N_1$ を $(N_1 - 1)\,K < N \leq N_1 K$ なる自然数とするとき，

$$J = \{0,\, K,\, 2K,\, \cdots,\, (N_1 - 1)K\}$$

に関するサブサンプラーを $(\downarrow K)$ と表す．

$\mathbf{e}_0 = (\delta_{n,0})_{n \in \boldsymbol{Z}_N} \in \boldsymbol{K}^N$ とする（$\delta_{n,0}$ はクロネッカーのデルタ記号である）．$\mathbf{e}_0$ を**単位インパルス**という．

**定理 19.3**　$J$ を (19.5) で定めたものとする．

$$\Delta_J = (\tau_{k_0}\mathbf{e}_0,\, \tau_{k_1}\mathbf{e}_0,\, \cdots,\, \tau_{k_{L-1}}\mathbf{e}_0) \in \boldsymbol{K}^{N \times L}$$

とおく．このとき $\mathbf{x} = (x[n])_{n \in \boldsymbol{Z}_N} \in \boldsymbol{K}^N$ に対して

$$\Delta_J^* \mathbf{x} = \Delta_J^T \mathbf{x} = (\downarrow J)\,\mathbf{x}$$

である．

**証明．** 定理 19.1 の証明より

$$\Delta_J^* \mathbf{x} = \begin{pmatrix} \mathbf{e}_0 \bigstar x[k_0] \\ \mathbf{e}_0 \bigstar x[k_1] \\ \vdots \\ \mathbf{e}_0 \bigstar x[k_{L-1}] \end{pmatrix} = \begin{pmatrix} x[k_0] \\ x[k_1] \\ \vdots \\ x[k_{L-1}] \end{pmatrix} = (\downarrow J)\,\mathbf{x}.$$

■

サブサンプラーを表す行列 $\Delta_J^*$ の共役行列 $(\Delta_J^*)^* = \Delta_J$ の表す操作を求める．

次の定義からはじめる．

**定義 19. 3**　(19.5) の $J$ と $\mathbf{y} = (y\,[k_j])_{j \in \boldsymbol{Z}_L} \in \boldsymbol{K}^L$ に対して

$$(\uparrow J)\ y\,[n] = \begin{cases} y\,[l]\,, & n = k_l \in J \\ 0, & n \in \boldsymbol{Z}_N,\ n \notin J \end{cases}$$

とし，

$$(\uparrow J)\,\mathbf{y} = ((\uparrow J)\ y\,[n])_{n \in \boldsymbol{Z}_N} \in \boldsymbol{K}^N$$

と定義する．$(\uparrow J)$ を $J$ に関する**アップサンプラー**という．

特に，$K$ を $K < N$ なる自然数で，

$$J = \{kK : k \in \boldsymbol{Z}_+,\ kK \leq N-1\}$$

とするときの $J$ に関するアップサンプラーを $(\uparrow K)$ により表す．

**定理 19. 4**　(19.5) の $J$ と $\mathbf{y} = (y\,[j])_{j \in \boldsymbol{Z}_L} \in \boldsymbol{K}^L$ に対して

$$\Delta_J\,\mathbf{y} = (\uparrow J)\,\mathbf{y}$$

である．

**証明．**

$$\Delta_J\,\mathbf{y} = (\tau_{k_0}\mathbf{e}_0,\ \tau_{k_1}\mathbf{e}_0,\ \cdots,\ \tau_{k_{L-1}}\mathbf{e}_0) \begin{pmatrix} y\,[0] \\ \vdots \\ y\,[L-1] \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{j=0}^{L-1} \tau_{k_j} e_0\,[0]\,y\,[j] \\ \displaystyle\sum_{j=0}^{L-1} \tau_{k_j} e_0\,[1]\,y\,[j] \\ \vdots \\ \displaystyle\sum_{j=0}^{L-1} \tau_{k_j} e_0\,[N-1]\,y\,[j] \end{pmatrix}.$$

ここで

$$\sum_{j=0}^{L-1} \tau_{k_j} e_0 [l] \, y [j] = \sum_{j=0}^{L-1} e_0 [l - k_j] \, y [j] = \begin{cases} y [n], & l = k_n \in J \\ 0, & l \notin J \end{cases}$$

$$= (\uparrow J) \, y [l] .$$

よって，$\Delta_J \mathbf{y} = (\uparrow J) \mathbf{y}$. ■

**命題 19.5**　行列 $A = (\mathbf{a}_0, \mathbf{a}_1, \cdots, \mathbf{a}_{N-1})$ に対して，

$$A\Delta_J = (\mathbf{a}_{k_0}, \mathbf{a}_{k_1}, \cdots, \mathbf{a}_{k_{L-1}}) \tag{19.6}$$

である．

**証明.** $A\Delta_J = \left( \sum_{\mu=0}^{N-1} a_{i\mu} \delta_{k_j} [\mu] \right) = (a_{ik_j}) = (\mathbf{a}_{k_0}, \mathbf{a}_{k_1}, \cdots, \mathbf{a}_{k_{L-1}})$. ■

**系 19.6**　$\mathbf{u} \in \boldsymbol{K}^N$ とし，

$$\mathcal{U} = (\tau_0 \mathbf{u}, \tau_1 \mathbf{u}, \cdots, \tau_{N-1} \mathbf{u}) \in \boldsymbol{K}^{N \times N},$$
$$\overline{\mathbf{u}} = (\overline{u [0]}, \cdots, \overline{u [N-1]})$$

とする．このとき $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$(\mathcal{U}\Delta_J)^* \mathbf{x} = (\downarrow J) (\overline{\mathbf{u}} \bigstar x)$$

である．また，$\mathbf{y} \in \boldsymbol{K}^L$ に対して，

$$\mathcal{U}\Delta_J \mathbf{y} = \mathbf{u} *_c (\uparrow J) \mathbf{y}$$

となっている．

## 19.2　サブバンド・フィルタ・バンク

ウェーブレット解析はスケーリング・フィルタとウェーブレット・フィルタの二つのフィルタによる処理であるが，これをより一般的な条件のもとで複数のフィルタによる処理が行えるように定式化する．

$n$ を自然数とする．$J_0, \cdots, J_{n-1} \subset \{0, 1, 2, \cdots, N-1\}$ とし，$J_j$ の元の個数を $L(j)$ とおく．

$$L = L(0) + \cdots + L(n-1)$$

とおく．$J_j$ に属する元を小さい順に並べたものを

$$k_{j,0},\ k_{j,1},\ k_{j,2},\ \cdots,\ k_{j,L(j)-1}$$

と表す．$J_j$ に関するサブサンプラーは，$\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$(\downarrow J_j)\,\mathbf{x} = (x\,[k_{j,l}])_{l \in \boldsymbol{Z}_{L(j)}} \in \boldsymbol{K}^{L(j)}$$

である．

**定義 19.4**　$\mathbf{u}_0, \cdots, \mathbf{u}_{n-1} \in \boldsymbol{K}^N$, $\widetilde{\mathbf{u}}_0, \cdots, \widetilde{\mathbf{u}}_{n-1} \in \boldsymbol{K}^N$ とする．

$$\begin{aligned}\mathcal{U}_\mu &= (\tau_0 \mathbf{u}_\mu,\, \tau_1 \mathbf{u}_\mu,\, \cdots,\, \tau_{N-1} \mathbf{u}_\mu) \in \boldsymbol{K}^{N \times N}, \quad \mu = 0, \cdots, n-1,\\ \widetilde{\mathcal{U}}_\mu &= (\tau_0 \widetilde{\mathbf{u}}_\mu,\, \tau_1 \widetilde{\mathbf{u}}_\mu,\, \cdots,\, \tau_{N-1} \widetilde{\mathbf{u}}_\mu) \in \boldsymbol{K}^{N \times N}, \quad \mu = 0, \cdots, n-1\end{aligned}$$

とおき，

$$\begin{aligned}U_\mu &= \mathcal{U}_\mu \Delta_{J_\mu} = \left(\tau_{k_{\mu,0}} \mathbf{u}_\mu,\, \cdots,\, \tau_{k_{\mu,L(\mu)-1}} \mathbf{u}_\mu\right) \in \boldsymbol{K}^{N \times L(\mu)}, \quad \mu = 0, \cdots, n-1,\\ \widetilde{U}_\mu &= \widetilde{\mathcal{U}}_\mu \Delta_{J_\mu} = \left(\tau_{k_{\mu,0}} \widetilde{\mathbf{u}}_\mu,\, \cdots,\, \tau_{k_{\mu,L(\mu)-1}} \widetilde{\mathbf{u}}_\mu\right) \in \boldsymbol{K}^{N \times L(\mu)}, \quad \mu = 0, \cdots, n-1\end{aligned}$$

とおく．また

$$W = \begin{pmatrix} U_0^* \\ \vdots \\ U_{n-1}^* \end{pmatrix} \in \boldsymbol{K}^{L \times N},$$

$$\widetilde{W} = \begin{pmatrix} \widetilde{U}_0 & \cdots & \widetilde{U}_{n-1} \end{pmatrix} \in \boldsymbol{K}^{N \times L}$$

とする．$\mathbf{u}_0, \cdots, \mathbf{u}_{n-1}$ を**分解フィルタ**，$\widetilde{\mathbf{u}}_0, \cdots, \widetilde{\mathbf{u}}_{n-1}$ を**合成フィルタ**，また $W$ を**分解バンク**，$\widetilde{W}$ を**合成バンク**という．そして

$$\text{組}\left\{W, \widetilde{W}\right\} \quad \text{あるいは} \quad \text{組}\,\{\mathbf{u}_0,\, \cdots,\, \mathbf{u}_{n-1}\,;\, \widetilde{\mathbf{u}}_0,\, \cdots,\, \widetilde{\mathbf{u}}_{n-1}\}$$

を本書では ($\boldsymbol{K}^N$ における) **フィルタ・バンク**という．

フィルタ・バンク $\{W, \widetilde{W}\}$ が

$$\widetilde{W} W = I_N$$

をみたすとき，このフィルタ・バンクは ($\boldsymbol{K}^N$ において) **完全再構成条件**をみたす，あるいは ($\boldsymbol{K}^N$ における) **完全再構成フィルタ・バンク**であるという．

特に $\{W, W^*\}$ が完全再構成条件をみたすとき，すなわち

$$W^* W = I_N$$

が成り立つとき，$W$ あるいは $\{W, W^*\}$ を本書では**パラユニタリ**であるという．

次にフィルタ・バンクに関するいくつかの有用なクラスを定義する．

**定義 19.5**　$\{W, \widetilde{W}\}$ を定義 19.4 のフィルタ・バンクとする．もし

$$J_0 = \cdots = J_{n-1}$$

が成り立つとき，$\{W, \widetilde{W}\}$ を**広義の $n$ チャネル等間引きフィルタ・バンク**という．

特に $N = M_1 N' \, (M_1, N' \in \boldsymbol{N})$ であり，かつ

$$J_0 = \cdots = J_{n-1} = \{0, M_1, 2M_1, \cdots, (N'-1)M_1\}$$

の場合，$\{W, \widetilde{W}\}$ を **$n$ チャネル $M_1$ 間引きフィルタ・バンク**という．

もし

$$L = N$$

が成り立つときは，$\{W, \widetilde{W}\}$ を**最大間引きフィルタ・バンク**という．また $L > N$ の場合は**過剰フィルタ・バンク**という．

$\{W, \widetilde{W}\}$ が完全再構成条件をみたすならば，$\widetilde{W}$ が $W$ の左逆行列なので，$\mathrm{rank}(W) = N$ となる必要がある (定理 15.5 参照)．同様にして $\mathrm{rank}(\widetilde{W}) = N$ であることも必要である．いずれにせよ，$L \geq N$ でなければならない．すなわち $L < N$ の場合には完全再構成条件はみたしえない．

**定理 19.7**　フィルタ・バンク $\{W, \widetilde{W}\}$ が完全再構成であるための必要十分条件は，次の (19.7) あるいは (19.8) が成り立つことである．

$$\sum_{\mu=0}^{n-1} \widetilde{U}_\mu U_\mu^* = I_N, \tag{19.7}$$

$$\sum_{\mu=0}^{n-1} U_\mu \widetilde{U}_\mu^* = I_N. \tag{19.8}$$

**証明.**

$$\widetilde{W} W = \begin{pmatrix} \widetilde{U}_0 & \cdots & \widetilde{U}_{n-1} \end{pmatrix} \begin{pmatrix} U_0^* \\ \vdots \\ U_{n-1}^* \end{pmatrix} = \sum_{\mu=0}^{n-1} \widetilde{U}_\mu U_\mu^*.$$

また

$$\left(\widetilde{W} W\right)^* = \sum_{\mu=0}^{n-1} U_\mu \widetilde{U}_\mu^*$$

であることから定理が証明される．■

この定理 19.7 は系 19.6 を用いて次のように言い換えることができる．

**定理 19.8**　フィルタ・バンク $\{W,\widetilde{W}\}$ が完全再構成であるための必要十分条件は，任意の $\mathbf{x}\in\boldsymbol{K}^N$ に対して

$$\mathbf{x}=\sum_{\mu=0}^{n}\widetilde{\mathbf{u}}_\mu *_c(\uparrow J_\mu)(\downarrow J_\mu)(\overline{\mathbf{u}}_\mu\bigstar\mathbf{x})$$

が成り立つことである．また，このことは任意の $\mathbf{x}\in\boldsymbol{K}^N$ に対して

$$\mathbf{x}=\sum_{\mu=0}^{n}\mathbf{u}_\mu *_c(\uparrow J_\mu)(\downarrow J_\mu)\left(\overline{\widetilde{\mathbf{u}}}_\mu\bigstar\mathbf{x}\right)$$

が成り立つこととも同値である．

**証明.**

$$\sum_{\mu=0}^{n-1}\widetilde{U}_\mu U_\mu^*\mathbf{x}=\sum_{\mu=0}^{n-1}\widetilde{\mathcal{U}}_\mu\Delta_{J_\mu}(\mathcal{U}_\mu\Delta_{J_\mu})^*\mathbf{x}=\sum_{\mu=0}^{n-1}\widetilde{\mathbf{u}}_\mu *_c(\uparrow J_\mu)(\downarrow J_\mu)(\overline{\mathbf{u}_\mu}\bigstar\mathbf{x})$$

であるから，前半の主張が証明される．後半の主張も同様である．■

特に $L=N$ の場合，すなわち最大間引きフィルタ・バンクの場合は次のことが成り立つ．

**定理 19.9**　最大間引きフィルタ・バンク $\{W,\widetilde{W}\}$ が完全再構成であるための必要十分条件は，$\mu,\nu\in\{0,\cdots,n-1\}$ に対して

$$U_\mu^*\widetilde{U}_\mu=I_{L(\mu)},$$
$$U_\mu^*\widetilde{U}_\nu=O_{L(\mu),L(\nu)},\ \mu\neq\nu$$

が成り立つことである．

**証明.** 最大間引きの場合，$W,\widetilde{W}$ は $N$ 次正方行列であるから，$\widetilde{W}W=I_N$ ならば $\widetilde{W}=W^{-1}$ となっている．したがって，

$$I_N=W\widetilde{W}=\begin{pmatrix}U_0^*\\ \vdots\\ U_{n-1}^*\end{pmatrix}\begin{pmatrix}\widetilde{U}_0 & \cdots & \widetilde{U}_{n-1}\end{pmatrix}$$

$$=\begin{pmatrix}U_0^*\widetilde{U}_0 & \cdots & U_0^*\widetilde{U}_{n-1}\\ \vdots & \ddots & \vdots\\ U_{n-1}^*\widetilde{U}_0 & \cdots & U_{n-1}^*\widetilde{U}_{n-1}\end{pmatrix}.$$

ゆえに定理の条件が成り立つ．逆もこの議論から示せる． ■

この定理は，$L = N$ の場合に $\widetilde{W}$ が $W$ の逆行列であり，$W\widetilde{W} = I_L$ も成り立っていることから証明された．しかし，$L > N$ の場合，$W$ の右逆行列は存在しないので，$W\widetilde{W} \neq I_L$ である．以下では $W\widetilde{W}$ について論ずる．

一般に $W\widetilde{W}$ は次のようになっている．

**命題 19.10**

$$W\widetilde{W} = \begin{pmatrix} U_0^*\widetilde{U}_0 & \cdots & U_0^*\widetilde{U}_{n-1} \\ \vdots & \ddots & \vdots \\ U_{n-1}^*\widetilde{U}_0 & \cdots & U_{n-1}^*\widetilde{U}_{n-1} \end{pmatrix}.$$

さらにこの行列の各ブロックは，次のようになっている．

$$U_\mu^*\widetilde{U}_\nu = \begin{pmatrix} \overline{u_\mu}\bigstar\widetilde{u}_\nu\left[k_{\mu,0} - k_{\nu,0}\right] & \cdots & \overline{u_\mu}\bigstar\widetilde{u}_\nu\left[k_{\mu,0} - k_{\nu,L(\nu)-1}\right] \\ \vdots & \ddots & \vdots \\ \overline{u_\mu}\bigstar\widetilde{u}_\nu\left[k_{\mu,L(\mu)-1} - k_{\nu,0}\right] & \cdots & \overline{u_\mu}\bigstar\widetilde{u}_\nu\left[k_{\mu,L(\mu)-1} - k_{\nu,L(\nu)-1}\right] \end{pmatrix}.$$

**証明.** ブロック行列の計算から

$$W\widetilde{W} = \begin{pmatrix} U_0^* \\ \vdots \\ U_{n-1}^* \end{pmatrix} \begin{pmatrix} \widetilde{U}_0 & \cdots & \widetilde{U}_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} U_0^*\widetilde{U}_0 & \cdots & U_0^*\widetilde{U}_{n-1} \\ \vdots & \ddots & \vdots \\ U_{n-1}^*\widetilde{U}_0 & \cdots & U_{n-1}^*\widetilde{U}_{n-1} \end{pmatrix}$$

である．各ブロックは次のようになっている．

$$U_\mu^*\widetilde{U}_\nu = \begin{pmatrix} (\tau_{k_{\mu,0}}\mathbf{u}_\mu)^* \\ \vdots \\ (\tau_{k_{\mu,L(\mu)-1}}\mathbf{u}_\mu)^* \end{pmatrix} \left(\tau_{k_{\nu,0}}\widetilde{\mathbf{u}}_\nu, \cdots, \tau_{k_{\nu,L(\nu)-1}}\widetilde{\mathbf{u}}_\nu\right)$$

$$
= \begin{pmatrix} \langle \tau_{k_{\nu,0}} \widetilde{\mathbf{u}}_\nu, \tau_{k_{\mu,0}} \mathbf{u}_\mu \rangle & \cdots & \left\langle \tau_{k_{\nu,L(\nu)-1}} \widetilde{\mathbf{u}}_\nu, \tau_{k_{\mu,0}} \mathbf{u}_\mu \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle \tau_{k_{\nu,0}} \widetilde{\mathbf{u}}_\nu, \tau_{k_{\mu,L(\mu)-1}} \mathbf{u}_\mu \right\rangle & \cdots & \left\langle \tau_{k_{\nu,L(\nu)-1}} \widetilde{\mathbf{u}}_\nu, \tau_{k_{\mu,L(\mu)-1}} \mathbf{u}_\mu \right\rangle \end{pmatrix}.
$$

ここで

$$
\begin{aligned}
\langle \tau_k \mathbf{v}, \tau_l \mathbf{w} \rangle &= \left\langle \begin{pmatrix} v[-k] \\ v[1-k] \\ \vdots \\ v[N-1-k] \end{pmatrix}, \begin{pmatrix} w[-l] \\ w[1-l] \\ \vdots \\ w[N-1-l] \end{pmatrix} \right\rangle \\
&= \sum_{r=0}^{N-1} v[r-k] \overline{w[r-l]} \\
&= \sum_{r=0}^{N-1} v[r] \overline{w[r+k-l]} \\
&= (v \bigstar \overline{w})[k-l] = (\overline{w} \bigstar v)[l-k]
\end{aligned}
$$

であるから，

$$
\begin{aligned}
& U_\mu^* \widetilde{U}_\nu \\
&= \begin{pmatrix} \overline{u_\mu} \bigstar \widetilde{u}_\nu [k_{\mu,0} - k_{\nu,0}] & \cdots & \overline{u_\mu} \bigstar \widetilde{u}_\nu [k_{\mu,0} - k_{\nu,L(\nu)-1}] \\ \vdots & \ddots & \vdots \\ \overline{u_\mu} \bigstar \widetilde{u}_\nu [k_{\mu,L(\mu)-1} - k_{\nu,0}] & \cdots & \overline{u_\mu} \bigstar \widetilde{u}_\nu [k_{\mu,L(\mu)-1} - k_{\nu,L(\nu)-1}] \end{pmatrix}.
\end{aligned}
$$

■

この命題から次のことが成り立つ．

**系 19.11** $\{W, \widetilde{W}\}$ が完全再構成フィルタ・バンクであるとする．$\widetilde{W}$ が $W$ のムーア-ペンローズ一般逆行列であるための必要十分条件は，$\mu, \nu \in \{0, \cdots, n-1\}$ に対して

$$
\widetilde{U}_\mu^* U_\nu = U_\mu^* \widetilde{U}_\nu
$$

が成り立つことである．特にパラユニタリフィルタ・バンク $\{W, W^*\}$ の場合，$W^*$ は $W$ のムーア-ペンローズ一般逆行列である．

**証明.** $\widetilde{W}$ が $W$ の左逆行列であることより，$W$ は $W$ の $\{1,2,4\}$-一般逆行列である．したがってムーア-ペンローズ一般逆行列であるための必要十分条件は，

$$\left(W\widetilde{W}\right)^* = W\widetilde{W}$$

である．

$$\left(W\widetilde{W}\right)^* = \begin{pmatrix} \widetilde{U}_0^* U_0 & \cdots & \widetilde{U}_0^* U_{n-1} \\ \vdots & \ddots & \vdots \\ \widetilde{U}_{n-1}^* U_0 & \cdots & \widetilde{U}_{n-1}^* U_{n-1} \end{pmatrix},$$

$$W\widetilde{W} = \begin{pmatrix} U_0^* \widetilde{U}_0 & \cdots & U_0^* \widetilde{U}_{n-1} \\ \vdots & \ddots & \vdots \\ U_{n-1}^* \widetilde{U}_0 & \cdots & U_{n-1}^* \widetilde{U}_{n-1} \end{pmatrix}$$

であることより，系を証明することができる．後半の主張は，$\widetilde{W} = W^*$ より

$$\left(W\widetilde{W}\right)^* = (WW^*)^* = WW^* = W\widetilde{W}$$

であることによる． ■

## 19.3　2 チャネル最大間引きフィルタ・バンク

$N = 2N_1$ とする．本節では応用上極めて重要な場合として，$n = 1$ で

$$J_0 = J_1 = \{0,\, 2,\, 4,\, \cdots,\, 2\,(N_1 - 1)\}$$

の場合を扱う．この場合のフィルタ・バンクを 2 チャネル最大間引きフィルタ・バンクという．特にこのフィルタ・バンクの完全再構成について論ずる．まず双直交性の定義から始める．

**定義 19.6**　$N, l$ を自然数とする．$\mathbf{a}_0, \cdots, \mathbf{a}_{l-1}, \mathbf{b}_0, \cdots, \mathbf{b}_{l-1} \in \boldsymbol{K}^N$ とし，

$$A = (\mathbf{a}_0, \cdots, \mathbf{a}_{l-1}), \quad B = (\mathbf{b}_0, \cdots, \mathbf{b}_{l-1})$$

とする．

$$A^* B = I_l$$

が成り立つとき，$\{\mathbf{a}_j\}_{j=0}^{l-1}$ は $\{\mathbf{b}_j\}_{j=0}^{l-1}$ に**双直交**であるという．

$A^* B = I_l$ ならば $B^* A = (A^* B)^* = I_l$ であるから，$\{\mathbf{a}_j\}_{j=0}^{l-1}$ が $\{\mathbf{b}_j\}_{j=0}^{l-1}$ に双直交ならば $\{\mathbf{b}_j\}_{j=0}^{l-1}$ は $\{\mathbf{a}_j\}_{j=0}^{l-1}$ に双直交である．このことから，$\{\mathbf{a}_j\}_{j=0}^{l-1}$ が $\{\mathbf{b}_j\}_{j=0}^{l-1}$ に双直交であることを $\{\mathbf{a}_j\}_{j=0}^{l-1}$ と $\{\mathbf{b}_j\}_{j=0}^{l-1}$ が**互いに双直交**であるともい

う．また

$$B^*A = \begin{pmatrix} \langle \mathbf{a}_0, \mathbf{b}_0 \rangle & \cdots & \langle \mathbf{a}_{l-1}, \mathbf{b}_0 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{a}_0, \mathbf{b}_{l-1} \rangle & \cdots & \langle \mathbf{a}_{l-1}, \mathbf{b}_{l-1} \rangle \end{pmatrix}$$

より，$\{\mathbf{a}_j\}_{j=0}^{l-1}$ と $\{\mathbf{b}_j\}_{j=0}^{l-1}$ が互いに**双直交であるための必要十分条件**は

$$\langle \mathbf{a}_\mu, \mathbf{b}_\nu \rangle = \begin{cases} 1, & \mu = \nu \\ 0, & \mu \neq \nu \end{cases}$$

である．特に $\{\mathbf{a}_j\}_{j=0}^{l-1}$ が正規直交系ならば，$\{\mathbf{a}_j\}_{j=0}^{l-1}$ は $\{\mathbf{a}_j\}_{j=0}^{l-1}$ に双直交である．

**定理 19. 12** $\mathbf{u}_0, \widetilde{\mathbf{u}}_0 \in \boldsymbol{K}^N$ で，$\{\tau_{2j}\mathbf{u}_0\}_{j=0}^{N_1-1}$ と $\{\tau_{2j}\widetilde{\mathbf{u}}_0\}_{j=0}^{N_1-1}$ が互いに双直交であるようなものが与えられたとする．このとき

$$u_1[k] = (-1)^k \overline{\widetilde{u}_0[1-k]},$$
$$\widetilde{u}_1[k] = (-1)^k \overline{u_0[1-k]}$$

$(k = 0, \cdots, N-1)$ により $\mathbf{u}_1$, $\widetilde{\mathbf{u}}_1 \in \boldsymbol{K}^N$ を定義する．$j = 0, 1$ に対して

$$U_j = (\tau_0 \mathbf{u}_j,\ \tau_2 \mathbf{u}_j,\ \cdots,\ \tau_{2(N_1-1)} \mathbf{u}_j) \in \boldsymbol{K}^{N \times N_1},$$
$$\widetilde{U}_j = (\tau_0 \widetilde{\mathbf{u}}_j, \tau_2 \widetilde{\mathbf{u}}_j,\ \cdots,\ \tau_{2(N_1-1)} \widetilde{\mathbf{u}}_j) \in \boldsymbol{K}^{N \times N_1}$$

とし，

$$W = \begin{pmatrix} U_0^* \\ U_1^* \end{pmatrix} \in \boldsymbol{K}^{N \times N},$$
$$\widetilde{W} = \begin{pmatrix} \widetilde{U}_0 & \widetilde{U}_1 \end{pmatrix} \in \boldsymbol{K}^{N \times N}$$

とする．このとき $\{W, \widetilde{W}\}$ は完全再構成フィルタ・バンクである．

**証明.** 双直交であることと命題 19.10 より，

$$I_{N_1} = U_0^* \widetilde{U}_0 = \begin{pmatrix} \overline{u_0} \bigstar \widetilde{u}_0\,[0] & \cdots & \overline{u_0} \bigstar \widetilde{u}_0\,[-2(N_1-1)] \\ \vdots & \ddots & \vdots \\ \overline{u_0} \bigstar \widetilde{u}_0\,[2(N_1-1)] & \cdots & \overline{u_0} \bigstar \widetilde{u}_0\,[0] \end{pmatrix}$$

である．また

$$\overline{u_1} \bigstar \widetilde{u}_1\,[k] = \sum_{n=0}^{N-1} \overline{u_1\,[n]}\, \widetilde{u}_1\,[n+k]$$

$$= \sum_{n=0}^{N-1} (-1)^n \widetilde{u}_0[1-n](-1)^{n+k}\overline{u_0[1-k-n]}$$

$$= (-1)^k \sum_{n=0}^{N-1} \overline{u_0[1-k-n]}\widetilde{u}_0[1-n]$$

$$= (-1)^k \sum_{n=0}^{N-1} \overline{u_0[n]}\widetilde{u}_0[n+k] = (-1)^k \overline{u_0} \bigstar \widetilde{u}_0\,[k].$$

ゆえに

$$U_1^* \widetilde{U}_1 = \begin{pmatrix} \overline{u_1}\bigstar\widetilde{u}_1\,[0] & \cdots & \overline{u_1}\bigstar\widetilde{u}_1\,[-2(N_1-1)] \\ \vdots & \ddots & \vdots \\ \overline{u_1}\bigstar\widetilde{u}_1\,[2(N_1-1)] & \cdots & \overline{u_1}\bigstar\widetilde{u}_1\,[0] \end{pmatrix}$$

$$= \begin{pmatrix} \overline{u_0}\bigstar\widetilde{u}_0\,[0] & \cdots & \overline{u_0}\bigstar\widetilde{u}_0\,[-2(N_1-1)] \\ \vdots & \ddots & \vdots \\ \overline{u_0}\bigstar\widetilde{u}_0\,[2(N_1-1)] & \cdots & \overline{u_0}\bigstar\widetilde{u}_0\,[0] \end{pmatrix}$$

$$= I_{N_1}.$$

さらに

$$U_1^* \widetilde{U}_0 = \begin{pmatrix} \overline{u_1}\bigstar\widetilde{u}_0\,[0] & \cdots & \overline{u_1}\bigstar\widetilde{u}_0\,[-2(N_1-1)] \\ \vdots & \ddots & \vdots \\ \overline{u_1}\bigstar\widetilde{u}_0\,[2(N_1-1)] & \cdots & \overline{u_1}\bigstar\widetilde{u}_0\,[0] \end{pmatrix}$$

であるが，ここで

$$\overline{u_1}\bigstar\widetilde{u}_0\,[k] = \sum_{n=0}^{N-1} \overline{u_1\,[n]}\widetilde{u_0}\,[n+k] = \sum_{n=0}^{N-1} (-1)^n \widetilde{u}_0[1-n]\widetilde{u_0}\,[n+k]$$

$$= -\sum_{n=0}^{N-1} (-1)^n \widetilde{u}_0[n]\widetilde{u_0}\,[1-n+k]$$

$$= -\sum_{m=0}^{N_1-1} \widetilde{u}_0[2m]\widetilde{u_0}\,[1-2m+k] + \sum_{m=0}^{N_1-1} \widetilde{u}_0[2m+1]\widetilde{u_0}\,[-2m+k].$$

$k = 2k'$ のときは，$-m + k' = m'$ とおくと $2m = 2k' - 2m'$ であるから

$$\sum_{m=0}^{N_1-1} \widetilde{u}_0[2m+1]\widetilde{u_0}\,[-2m+2k'] = \sum_{m=0}^{N_1-1} \widetilde{u}_0[1-2m+2k']\widetilde{u_0}\,[2m].$$

したがって,

$$\overline{u_1}\bigstar\widetilde{u}_0\,[2k'] = 0$$

であることがわかる．ゆえに $U_1^*\widetilde{U}_0 = O_{N_1}$ である．同様にして $U_0^*\widetilde{U}_1 = O_{N_1}$ も示せる．よって定理 19.9 より定理が証明される． ■

**系 19.13** $N = 2N_1$ とする．$\mathbf{u}_0 \in \boldsymbol{K}^N$ であり，$\{\tau_{2j}\mathbf{u}_0\}_{j=0}^{N_1-1}$ が正規直交系であるとする．このとき

$$u_1[k] = (-1)^k\overline{u_0[1-k]}$$

$(k = 0, \cdots, N-1)$ により $\mathbf{u}_1 \in \boldsymbol{K}^N$ を定義し，$j = 0, 1$ に対して

$$U_j = (\tau_0\mathbf{u}_j,\, \tau_2\mathbf{u}_j,\, \cdots,\, \tau_{2(N_1-1)}\mathbf{u}_j) \in \boldsymbol{K}^{N\times N_1}$$

とし,

$$W = \begin{pmatrix} U_0^* \\ U_1^* \end{pmatrix} \in \boldsymbol{K}^{N\times N}$$

とすると，$\{W, W^*\}$ はパラユニタリフィルタ・バンクである．

さて，一般に

$$\operatorname{span}\{\tau_{2j}\mathbf{u}\}_{j=0}^{N_1-1} = \operatorname{span}\{\tau_{2j}\widetilde{\mathbf{u}}\}_{j=0}^{N_1-1}, \tag{19.9}$$

$$\operatorname{span}\{\tau_{2j}\mathbf{u}\}_{j=0}^{N_1-1} \neq \operatorname{span}\{\tau_{2j}\widetilde{\mathbf{u}}\}_{j=0}^{N_1-1} \tag{19.10}$$

の二つの場合が考えられる．たとえば第 21.1 節で学ぶ双対基底を用いて (19.9) をみたすような基底 $\{\tau_{2j}\widetilde{\mathbf{u}}\}_{j=0}^{N_1-1}$ の存在を証明することができるが，実際に利用されているものとしては (19.10) をみたすものが多い．

**問題 19.2** $\mathbf{u}, \widetilde{\mathbf{u}} \in \boldsymbol{K}^8$ を

$$u[k] = \begin{cases} \dfrac{1}{\sqrt{2}} & k = 0, 1 \\ 0 & k = 2, \cdots, 7 \end{cases}, \quad \widetilde{u}[k] = \begin{cases} -\dfrac{1}{8\sqrt{2}} & k = 6, 3 \\ \dfrac{1}{8\sqrt{2}} & k = 7, 2 \\ \dfrac{1}{\sqrt{2}} & k = 0, 1 \\ 0 & k = 4, 5 \end{cases}$$

により定義する．このとき，$U^*\widetilde{U} = I_4$ であること，すなわち $\mathbf{u}, \widetilde{\mathbf{u}}$ が互いに双直交であることを示せ．また，$\operatorname{span}\{\tau_{2j}\mathbf{u}\}_{j=0}^{3} \neq \operatorname{span}\{\tau_{2j}\widetilde{\mathbf{u}}\}_{j=0}^{3}$ であることを示

せ(このフィルタは [CDF] の方法により作られる．[CDF] では応用性の高い双直交基底と最大間引き完全再構成フィルタの構成方法が記されている)．

## 19.4　多重解像度近似

本節ではウェーブレットで重要な役割を果たす多重解像度分解の概念を述べる．

### 19.4.1　一般化多重解像度近似

$p, N_0, M_1, \cdots, M_p, N_1, \cdots, N_p$ を自然数で，

$$N_0 = M_1 N_1, \quad N_1 = M_2 N_2, \quad \cdots, \quad N_{p-1} = M_p N_p$$

をみたすものとする．便宜上 $N_0 = N$ とも表す．

$$\left\{\mathbf{u}_0^{(1)}, \cdots, \mathbf{u}_{m_1-1}^{(1)}; \widetilde{\mathbf{u}}_0^{(1)}, \cdots, \widetilde{\mathbf{u}}_{m_1-1}^{(1)}\right\}$$

を $\boldsymbol{K}^{N_0}$ における完全再構成な $m_1$ チャネル $M_1$ 間引きフィルタ・バンクとする．また

$$\left\{\mathbf{u}_0^{(2)}, \cdots, \mathbf{u}_{m_2-1}^{(2)}; \widetilde{\mathbf{u}}_0^{(2)}, \cdots, \widetilde{\mathbf{u}}_{m_2-1}^{(2)}\right\}$$

を $\boldsymbol{K}^{N_1}$ における完全再構成 $m_2$ チャネル $M_2$ 間引きフィルタ・バンク，以下同様に

$$\left\{\mathbf{u}_0^{(j)}, \cdots, \mathbf{u}_{m_j-1}^{(j)}; \widetilde{\mathbf{u}}_0^{(j)}, \cdots, \widetilde{\mathbf{u}}_{m_j-1}^{(j)}\right\}$$

を $\boldsymbol{K}^{N_{j-1}}$ における完全再構成 $m_j$ チャネル $M_j$ 間引きフィルタ・バンク ($j = 3, \cdots, p$) とする．

本書では $\mathbf{u}_0^{(j)}, \widetilde{\mathbf{u}}_0^{(j)}$ として低域通過フィルタをとり，

$$\mathbf{u}_1^{(j)}, \cdots, \mathbf{u}_{m_j-1}^{(j)}, \widetilde{\mathbf{u}}_1^{(j)}, \cdots, \widetilde{\mathbf{u}}_{m_j-1}^{(j)}, \quad j = 1, \cdots, p$$

は帯域通過フィルタをとることを想定している．各 $m_j$ チャネル $M_j$ 間引きフィルタ・バンクに対して，定義 19.4 と同様に定義した分解バンクと合成バンクを

$$W^{(j)} = \begin{pmatrix} U_0^{(j)*} \\ \vdots \\ U_{m_j-1}^{(j)*} \end{pmatrix}, \quad \widetilde{W}^{(j)} = \begin{pmatrix} \widetilde{U}_0^{(j)} & \cdots & \widetilde{U}_{m_j-1}^{(j)} \end{pmatrix}$$

とする．ただしここで

$$U_\mu^{(j)} \in \boldsymbol{K}^{N_{j-1} \times N_j}, \ \widetilde{U}_\mu^{(j)} \in \boldsymbol{K}^{N_{j-1} \times N_j}, \quad j = 1, \cdots, p, \ \mu = 0, \cdots, m_j - 1$$

である．

以下で用いるフィルタ・バンクに関する記号は定義 19.4 と同様にして定めるものとする．

$\mathbf{x} \in \boldsymbol{K}^{N_0}$ とする．

$$\mathbf{y}_\mu^{(1)} = U_\mu^{(1)*}\mathbf{x} \in \boldsymbol{K}^{N_1}, \quad \mu = 0, \cdots, m_1 - 1$$

をレベル 1 の**サブバンド信号**という．通常信号の低域部分は $\mathbf{y}_0^{(1)}$ に集中していることが多い．そこで低域部分 $\mathbf{y}_0^{(1)}$ をさらに分解する．

$$\mathbf{y}_\mu^{(2)} = U_\mu^{(2)*}\mathbf{y}_0^{(1)} \in \boldsymbol{K}^{N_2}, \quad \mu = 0, \cdots, m_2 - 1$$

とする．そして

$$\mathbf{y}_0^{(2)},\ \mathbf{y}_1^{(2)},\ \cdots, \mathbf{y}_{m_2-1}^{(2)},\ \mathbf{y}_1^{(1)},\ \cdots,\ \mathbf{y}_{m_1-1}^{(1)}$$

をレベル 2 のサブバンド信号という．

もしここで現れた低域部分 $\mathbf{y}_0^{(2)}$ にまだ除きたい高周波部分が含まれている場合は，$\mathbf{y}_0^{(2)}$ に対して，$\boldsymbol{K}^{N_2}$ における完全再構成 $m_3$ チャネル $M_3$ 間引きフィルタ・バンク $\left\{\mathbf{u}_0^{(3)}, \cdots, \mathbf{u}_{m_3-1}^{(3)}; \widetilde{\mathbf{u}}_0^{(3)}, \cdots, \widetilde{\mathbf{u}}_{m_3-1}^{(3)}\right\}$ を用いて同様の分解をする．必要ならこの操作を続けていく．このようにして次のようなベクトルが得られる．

$$\begin{aligned}
\mathbf{y}_\mu^{(1)} &= U_\mu^{(1)*}\mathbf{x} \in K^{N_1}, \quad \mu = 1, \cdots, m_1 - 1, \\
\mathbf{y}_\mu^{(2)} &= U_\mu^{(2)*}U_0^{(1)*}\mathbf{x} \in K^{N_2}, \quad \mu = 1, \cdots, m_2 - 1, \\
\mathbf{y}_\mu^{(3)} &= U_\mu^{(3)*}U_0^{(2)*}U_0^{(1)*}\mathbf{x} \in K^{N_3}, \quad \mu = 1, \cdots, m_3 - 1, \\
&\ \vdots \\
\mathbf{y}_\mu^{(p)} &= U_\mu^{(p)*}U_0^{(p-1)*} \cdots U_0^{(1)*}\mathbf{x} \in K^{N_p}, \quad \mu = 1, \cdots, m_p - 1, \\
\mathbf{y}_0^{(p)} &= U_0^{(p)*}U_0^{(p-1)*} \cdots U_0^{(1)*}\mathbf{x} \in K^{N_p}.
\end{aligned}$$

これをレベル $p$ のサブバンド信号という．このサブバンド信号に何らかの処理あるいは符号化が施されることになる．

次にサブバンド信号からもとの信号 $\mathbf{x}$ を復元する操作を述べる．そのために $\widetilde{\mathbf{u}}_\mu^{(k)}$ を用いて

$$\begin{aligned}
\mathbf{x}_{p,0} &= \widetilde{U}_0^{(1)}\widetilde{U}_0^{(2)} \cdots \widetilde{U}_0^{(p)}U_0^{(p)*}U_0^{(p-1)*} \cdots U_0^{(1)*}\mathbf{x} \in K^N, \\
\mathbf{x}_{p,\mu} &= \widetilde{U}_0^{(1)}\widetilde{U}_0^{(2)} \cdots \widetilde{U}_\mu^{(p)}U_\mu^{(p)*}U_0^{(p-1)*} \cdots U_0^{(1)*}\mathbf{x} \in K^N,
\end{aligned}$$

$$
\begin{aligned}
\mathbf{x}_{p-1,\mu} &= \widetilde{U}_0^{(1)}\widetilde{U}_0^{(2)}\cdots\widetilde{U}_\mu^{(p-1)}U_\mu^{(p-1)*}U_0^{(p-2)*}\cdots U_0^{(1)*}\mathbf{x} \in K^N, \\
&\ \vdots \\
\mathbf{x}_{2,\mu} &= \widetilde{U}_0^{(1)}\widetilde{U}_\mu^{(2)}U_\mu^{(2)*}U_0^{(1)*}\mathbf{x} \in K^N, \\
\mathbf{x}_{1,\mu} &= \widetilde{U}_\mu^{(1)}U_\mu^{(1)*}\mathbf{x} \in K^N
\end{aligned}
$$

とおく．このとき完全再構成条件 (定理 19.7 参照) から

$$
\begin{aligned}
&\mathbf{x}_{p,0} + \sum_{\mu=1}^{m_p-1}\mathbf{x}_{p,\mu} + \sum_{\mu=1}^{m_{p-1}-1}\mathbf{x}_{p-1,\mu} + \cdots + \sum_{\mu=1}^{m_1-1}\mathbf{x}_{1,\mu} \\
&= \widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(p-1)}\left(\sum_{\mu=0}^{m_p-1}\widetilde{U}_\mu^{(p)}U_\mu^{(p)*}\right)U_0^{(p-1)*}\cdots U_0^{(1)*}\mathbf{x} \\
&\quad + \sum_{\mu=1}^{m_{p-1}-1}\mathbf{x}_{p-1,\mu} + \cdots + \sum_{\mu=1}^{m_1-1}\mathbf{x}_{1,\mu} \\
&= \widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(p-1)}U_0^{(p-1)*}\cdots U_0^{(1)*}\mathbf{x} + \sum_{\mu=1}^{m_{p-1}-1}\mathbf{x}_{p-1,\mu} + \cdots + \sum_{\mu=1}^{m_1-1}\mathbf{x}_{1,\mu} \\
&= \widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(p-2)}\left(\sum_{\mu=0}^{m_{p-1}-1}\widetilde{U}_\mu^{(p-1)}U_\mu^{(p-1)*}\right)U_0^{(p-2)*}\cdots U_0^{(1)*}\mathbf{x} \\
&\quad + \sum_{\mu=1}^{m_{p-2}-1}\mathbf{x}_{p-2,\mu} + \cdots + \sum_{\mu=1}^{m_1-1}\mathbf{x}_{1,\mu} \\
&= \widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(p-2)}U_0^{(p-2)*}\cdots U_0^{(1)*}\mathbf{x} + \sum_{\mu=1}^{m_{p-2}-1}\mathbf{x}_{p-2,\mu} + \cdots + \sum_{\mu=1}^{m_1-1}\mathbf{x}_{1,\mu} \\
&= \cdots = \sum_{\mu=0}^{m_1-1}\widetilde{U}_\mu^{(1)}U_\mu^{(1)*}\mathbf{x} = \mathbf{x}.
\end{aligned}
$$

そこで

$$
\begin{aligned}
\widetilde{P}_\mu^{(1)} &= \widetilde{U}_\mu^{(1)}U_\mu^{(1)*}, \\
\widetilde{P}_\mu^{(2)} &= \widetilde{U}_0^{(1)}\widetilde{U}_\mu^{(2)}U_\mu^{(2)*}U_0^{(1)*}, \\
&\ \vdots
\end{aligned}
$$

$$\widetilde{P}_\mu^{(p)} = \widetilde{U}_0^{(1)}\widetilde{U}_0^{(2)}\cdots\widetilde{U}_\mu^{(p)}U_\mu^{(p)*}U_0^{(p-1)*}\cdots U_0^{(1)*},$$

$$\widetilde{P}_0^{(p)} = \widetilde{U}_0^{(1)}\widetilde{U}_0^{(2)}\cdots\widetilde{U}_0^{(p)}U_0^{(p)*}U_0^{(p-1)*}\cdots U_0^{(1)*}$$

とおくと，

$$I_N = \widetilde{P}_0^{(p)} + \sum_{\mu=1}^{m_p-1}\widetilde{P}_\mu^{(p)} + \sum_{\mu=1}^{m_{p-1}-1}\widetilde{P}_\mu^{(p-1)} + \cdots + \sum_{\mu=1}^{m_1-1}\widetilde{P}_\mu^{(1)}$$

が成り立つ．この分解を**多重解像度分解**という．

$\mathbf{x} \in K^N$ に対して $\widetilde{P}_0^{(p)}\mathbf{x}$ を $\mathbf{x}$ のレベル $p$ の**近似部分**，$j = 1, 2, \cdots, p$ に対して $\sum_{\mu=1}^{m_j-1}\widetilde{P}_\mu^{(j)}\mathbf{x}$ を $\mathbf{x}$ のレベル $j$ の**詳細部分**という．

**注意 19.1** $\mathbf{u}_\mu^{(j)}$ と $\widetilde{\mathbf{u}}_\mu^{(j)}$ の役割を入れ替えて，本節と同じ議論をすることも可能である．この場合，

$$P_\mu^{(1)} = U_\mu^{(1)}\widetilde{U}_\mu^{(1)*},$$

$$P_\mu^{(2)} = U_0^{(1)}U_\mu^{(2)}\widetilde{U}_\mu^{(2)*}\widetilde{U}_0^{(1)*},$$

$$\vdots$$

$$P_\mu^{(p)} = U_0^{(1)}U_0^{(2)}\cdots U_\mu^{(p)}\widetilde{U}_\mu^{(p)*}\widetilde{U}_0^{(p-1)*}\cdots\widetilde{U}_0^{(1)*}$$

とすると，

$$I_N = P_0^{(p)} + \sum_{\mu=1}^{m_p-1}P_\mu^{(p)} + \sum_{\mu=1}^{m_{p-1}-1}P_\mu^{(p-1)} + \cdots + \sum_{\mu=1}^{m_1-1}P_\mu^{(1)}$$

である．

### 19.4.2 最大間引きフィルタ・バンクと多重解像度解析

第 19.4.1 節において特に各 $\left\{\mathbf{u}_0^{(j)}, \cdots, \mathbf{u}_{m_j-1}^{(j)}; \widetilde{\mathbf{u}}_0^{(j)}, \cdots, \widetilde{\mathbf{u}}_{m_j-1}^{(j)}\right\}$ が $\boldsymbol{K}^{N_{j-1}}$ の $M_j$ チャネル最大間引き完全再構成フィルタ・バンクをなしている場合を扱う ($j = 1, \cdots, p$，ただし，便宜上 $N_0 = N$ とおく．またこの仮定から $m_j = M_j$ となっている)．このときは，多重解像度近似は単に近似するだけでなく，$\boldsymbol{K}^N$ の部分空間による直和分解も与えている．このことについて述べる．

はじめに最大間引きの場合，$\widetilde{P}_\mu^{(j)}$ は射影行列になっていることを示しておく．実際，定理 19.9 より

$$U_0^{(k)*}\widetilde{U}_0^{(k)} = I_{N_k}, \quad k = 1, \cdots, p$$

であるから

$$\begin{aligned}\widetilde{P}_{\mu}^{(j)}\widetilde{P}_{\mu}^{(j)} &= \left(\widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(j-1)}\widetilde{U}_{\mu}^{(j)}U_{\mu}^{(j)*}U_0^{(j-1)*}\cdots U_0^{(1)*}\right)\\ &\quad\times\left(\widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(j-1)}\widetilde{U}_{\mu}^{(j)}U_{\mu}^{(j)*}U_0^{(j-1)*}\cdots U_0^{(1)*}\right)\\ &= \left(\widetilde{U}_0^{(1)}\cdots\widetilde{U}_0^{(j-1)}\widetilde{U}_{\mu}^{(j)}\right)\left(U_{\mu}^{(j)*}U_0^{(j-1)*}\cdots U_0^{(1)*}\right)\\ &= \widetilde{P}_{\mu}^{(j)}\end{aligned}$$

(ただしここで $\times$ は行列の積を表すものとする) となっている．同様に $P_{\mu}^{(j)}P_{\mu}^{(j)} = P_{\mu}^{(j)}$ も示せる．

**レベル 1 の多重解像度解析**

$\mu \neq \nu$ のとき，定理 19.9 より

$$\widetilde{P}_{\mu}^{(1)}\widetilde{P}_{\nu}^{(1)} = \widetilde{U}_{\mu}^{(1)}U_{\mu}^{(1)*}\widetilde{U}_{\nu}^{(1)}U_{\nu}^{(1)*} = \widetilde{U}_{\mu}^{(1)}O_{N_1,N_1}U_{\nu}^{(1)*} = O_{N,N}$$

である．

$$\begin{aligned}\widetilde{V}_1 &= \mathcal{R}\left(\widetilde{P}_0^{(1)}\right),\\ \widetilde{W}_{1,\mu} &= \mathcal{R}\left(\widetilde{P}_{\mu}^{(1)}\right),\quad \mu = 1,\cdots,M_1-1\end{aligned}$$

とおくと，次の直和分解が成り立っている．

$$\boldsymbol{K}^N = \widetilde{V}_1 \dotplus \widetilde{W}_{1,1} \dotplus \cdots \dotplus \widetilde{W}_{1,M_1-1}. \tag{19.11}$$

**証明.** $I_N = \sum_{\mu=0}^{M_1-1}\widetilde{P}_{\mu}^{(1)}$ より任意の $\mathbf{x} \in \boldsymbol{K}^N$ は $\mathbf{x} = \sum_{\mu=0}^{M_1-1}\widetilde{P}_{\mu}^{(1)}\mathbf{x}$ と分解され，さらに $\mu \neq \nu$ のとき，$\mathbf{y} \in \widetilde{W}_{1,\mu} \cap \widetilde{W}_{1,\nu}$ であれば，

$$\mathbf{y} = \widetilde{P}_{\mu}^{(1)}\mathbf{x} = \widetilde{P}_{\nu}^{(1)}\mathbf{x}'$$

より

$$\mathbf{y} = \widetilde{P}_{\mu}^{(1)}\mathbf{x} = \widetilde{P}_{\mu}^{(1)}\widetilde{P}_{\mu}^{(1)}\mathbf{x} = \widetilde{P}_{\mu}^{(1)}\widetilde{P}_{\nu}^{(1)}\mathbf{x}' = \mathbf{0}$$

となっている．ゆえに $\widetilde{W}_{1,\mu} \cap \widetilde{W}_{1,\nu} = \{\mathbf{0}\}$ である．同様にして $\widetilde{V}_1 \cap \widetilde{W}_{1,\mu} = \{\mathbf{0}\}$ である． ■

ここで $\widetilde{V}_1$ は低域にエネルギーの集中したベクトルからなっていて，各 $\widetilde{W}_{1,\mu}$, $\mu > 0$ は比較的高い帯域にエネルギーの集中したベクトルからなっていると考えることができる．そこで高域の部分をまとめて

$$\widetilde{W}_1 = \widetilde{W}_{1,1} \dotplus \cdots \dotplus \widetilde{W}_{1,M_1-1}$$

とおけば,

$$\boldsymbol{K}^N = \widetilde{V}_1 \dotplus \widetilde{W}_1 \tag{19.12}$$

が成り立つ.

分解 (19.11) あるいは (19.12) を**レベル 1 の多重解像度解析**という.

**レベル 2 の多重解像度解析**

$\mu \in \boldsymbol{Z}_{M_1}$, $\nu \in \boldsymbol{Z}_{M_2}$, $\mu \neq 0$ に対して,

$$\begin{aligned}\widetilde{P}_\mu^{(1)} \widetilde{P}_\nu^{(2)} &= \widetilde{U}_\mu^{(1)} \left( U_\mu^{(1)*} \widetilde{U}_0^{(1)} \right) \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} U_0^{(1)*} \\ &= \widetilde{U}_\mu^{(1)} O \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} U_0^{(1)*} = O, \\ \widetilde{P}_\nu^{(2)} \widetilde{P}_\mu^{(1)} &= \widetilde{U}_0^{(1)} \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} \left( U_0^{(1)*} \widetilde{U}_\mu^{(1)} \right) U_\mu^{(1)*} \\ &= \widetilde{U}_0^{(1)} \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} O U_\mu^{(1)*} = O\end{aligned}$$

であり, また $\nu, \mu \in \boldsymbol{Z}_{M_2}$, $\mu \neq \nu$ に対して,

$$\begin{aligned}\widetilde{P}_\mu^{(2)} \widetilde{P}_\nu^{(2)} &= \widetilde{U}_0^{(1)} \widetilde{U}_\mu^{(2)} U_\mu^{(2)*} U_0^{(1)*} \widetilde{U}_0^{(1)} \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} U_0^{(1)*} \\ &= \widetilde{U}_0^{(1)} \widetilde{U}_\mu^{(2)} U_\mu^{(2)*} \widetilde{U}_\nu^{(2)} U_\nu^{(2)*} U_0^{(1)*} \\ &= O \end{aligned} \tag{19.13}$$

が成り立っている.

$$\widetilde{V}_2 = \mathcal{R}\left(\widetilde{P}_0^{(2)}\right),$$

$$\widetilde{W}_{2,\mu} = \mathcal{R}\left(\widetilde{P}_\mu^{(2)}\right), \quad \mu = 1, \cdots, M_2 - 1$$

とおく. このとき

$$\widetilde{V}_1 = \widetilde{V}_2 \dotplus \widetilde{W}_{2,1} \dotplus \cdots \dotplus \widetilde{W}_{2,M_2-1} \tag{19.14}$$

となっていることを証明する.

**証明.** まず $\mu \in \{0, 1, \cdots, M_2 - 1\}$ に対して, $\mathcal{R}\left(\widetilde{P}_\mu^{(2)}\right) \subset \mathcal{R}\left(\widetilde{P}_0^{(1)}\right) = \widetilde{V}_1$ を示す. 完全再構成条件より

$$\widetilde{P}_\mu^{(2)} = I_N \widetilde{P}_\mu^{(2)} = \sum_{\nu=0}^{M_1-1} \widetilde{P}_\nu^{(1)} \widetilde{P}_\mu^{(2)} = \sum_{\substack{\nu=0 \\ \nu \neq \mu}}^{M_1-1} \widetilde{P}_\nu^{(1)} \widetilde{P}_\mu^{(2)} + \widetilde{P}_\mu^{(1)} \widetilde{P}_\mu^{(2)}$$

であるから，$\mu \neq 0$ の場合，

$$\widetilde{P}_\mu^{(2)} = \sum_{\substack{\nu=0 \\ \nu\neq\mu}}^{M_1-1} \widetilde{P}_\nu^{(1)} \widetilde{P}_\mu^{(2)} = \sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{M_1-1} \widetilde{P}_\nu^{(1)} \widetilde{P}_\mu^{(2)} + \widetilde{P}_0^{(1)} \widetilde{P}_\mu^{(2)} = \widetilde{P}_0^{(1)} \widetilde{P}_\mu^{(2)}$$

である．また $\mu = 0$ の場合は

$$\widetilde{P}_\mu^{(2)} = \sum_{\nu=1}^{M_1-1} \widetilde{P}_\nu^{(1)} \widetilde{P}_0^{(2)} + \widetilde{P}_0^{(1)} \widetilde{P}_0^{(2)} = \widetilde{P}_0^{(1)} \widetilde{P}_0^{(2)}$$

となっている．したがって，$\mathcal{R}\left(\widetilde{P}_\mu^{(2)}\right) = \mathcal{R}\left(\widetilde{P}_0^{(1)} \widetilde{P}_\mu^{(2)}\right) \subset \mathcal{R}\left(\widetilde{P}_0^{(1)}\right) = \widetilde{V}_1$ が成り立っている．ゆえに $\widetilde{V}_2, \widetilde{W}_{2,1}, \cdots, \widetilde{W}_{2,M_2-1}$ は $\widetilde{V}_1$ の部分空間である．さらに

$$\sum_{\mu=0}^{M_2-1} \widetilde{P}_\mu^{(2)} = \sum_{\mu=0}^{M_2-1} \widetilde{U}_0^{(1)} \widetilde{U}_\mu^{(2)} U_\mu^{(2)*} U_0^{(1)*} = \widetilde{U}_0^{(1)} U_0^{(1)*} = \widetilde{P}_0^{(1)}$$

であるから，(19.13) と併せて，分解 (19.14) を証明することができる．■

(19.14) と (19.12) より

$$\begin{aligned} \boldsymbol{K}^N &= \widetilde{V}_1 \dotplus \widetilde{W}_1 \\ &= \widetilde{V}_2 \dotplus \widetilde{W}_{2,1} \dotplus \cdots \dotplus \widetilde{W}_{2,M_2-1} \dotplus \widetilde{W}_1 \end{aligned} \tag{19.15}$$

が成り立つ．ここで

$$\widetilde{W}_2 = \widetilde{W}_{2,1} \dotplus \cdots \dotplus \widetilde{W}_{2,M_2-1}$$

とおくと

$$\boldsymbol{K}^N = \widetilde{V}_2 \dotplus \widetilde{W}_2 \dotplus \widetilde{W}_1 \tag{19.16}$$

が成り立つ．分解 (19.15) あるいは (19.16) を**レベル 2 の多重解像度解析**という．

**レベル $p$ の多重解像度解析**

以上の考察と同様にして，$j = 3, \cdots, p$ に対して

$$\begin{aligned} \widetilde{V}_j &= \mathcal{R}\left(\widetilde{P}_0^{(j)}\right), \\ \widetilde{W}_{j,\mu} &= \mathcal{R}\left(\widetilde{P}_\mu^{(j)}\right), \quad \mu = 1, \cdots, M_j - 1 \end{aligned}$$

とすると，次の直和分解が得られる．

$$\boldsymbol{K}^N = \widetilde{V}_p \dotplus \widetilde{W}_p \dotplus \widetilde{W}_{p-1} \dotplus \cdots \dotplus \widetilde{W}_1. \tag{19.17}$$

ただし，ここで

$$\widetilde{W}_j = \widetilde{W}_{j,1} \dotplus \cdots \dotplus \widetilde{W}_{j,M_j-1} \tag{19.18}$$

である．この分解を**レベル $p$ の多重解像度解析**という．

ところで，

$$V_j = \mathcal{R}\left(P_0^{(j)}\right),$$
$$W_{j,\mu} = \mathcal{R}\left(P_\mu^{(j)}\right), \quad \mu = 1, \cdots, M_j - 1$$

とすると，同様にして次の直和分解も得られる．

$$\boldsymbol{K}^N = V_p \dotplus W_p \dotplus W_{p-1} \dotplus \cdots \dotplus W_1. \tag{19.19}$$

ただし，

$$W_j = W_{j,1} \dotplus \cdots \dotplus W_{j,M_j-1}. \tag{19.20}$$

この分解も**レベル $p$ の多重解像度解析**という．

特に最大間引きパラユニタリ・フィルタ・バンクの場合は $V_j = \widetilde{V}_j$, $W_{j,\mu} = \widetilde{W}_{j,\mu}$ であり，さらに多重解像度解析における直和分解は直交分解になっている．

**問題 19. 3** $\boldsymbol{K}^N$ における 2 分割最大間引きフィルタ・バンク $\{\mathbf{u}_0, \mathbf{u}_1; \widetilde{\mathbf{u}}_0, \widetilde{\mathbf{u}}_1\}$ を考える．任意の $\mathbf{x} \in \boldsymbol{K}^N$ に対して

$$P_0\mathbf{x} = \sum_{j=0}^{M-1} (\mathbf{x}, \tau_{2j}\widetilde{\mathbf{u}}_0)\, \tau_{2j}\mathbf{u}_0, \qquad P_1\mathbf{x} = \sum_{j=0}^{M-1} (\mathbf{x}, \tau_{2j}\widetilde{\mathbf{u}}_1)\, \tau_{2j}\mathbf{u}_1,$$

$$\widetilde{P}_0\mathbf{x} = \sum_{j=0}^{M-1} (\mathbf{x}, \tau_{2j}\mathbf{u}_0)\, \tau_{2j}\widetilde{\mathbf{u}}_0, \qquad \widetilde{P}_1\mathbf{x} = \sum_{j=0}^{M-1} (\mathbf{x}, \tau_{2j}\mathbf{u}_1)\, \tau_{2j}\widetilde{\mathbf{u}}_1$$

であることを示せ．

## 19. 5 ウェーブレットの例

### 19. 5. 1 ハール・ウェーブレット

ハール・ウェーブレットはウェーブレットの中で最も単純なものである．これは次の二つのベクトルからなる．

$$\mathbf{u}_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\\vdots\\0\end{pmatrix}, \quad \mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\\vdots\\0\end{pmatrix} \in \boldsymbol{K}^{N_0}.$$

このうち，(19.3), (19.4) より $\mathbf{u}_0$ がスケーリング・フィルタであり，$\mathbf{u}_1$ がウェーブレット・フィルタであることが容易に確認できる．$\mathbf{u}_0$ をハール・スケーリング・フィルタ，$\mathbf{u}_1$ をハール・ウェーブレット・フィルタという．これをもとに 2 分割最大間引きフィルタ・バンクを構成する．$N_0 = 2N_1$ であるとする．

$$U_0^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \in \boldsymbol{K}^{N_1 \times N_0},$$

$$U_1^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\in \boldsymbol{K}^{N_1 \times N_0}$$

であり，

$$W = \begin{pmatrix} U_0^* \\ U_1^* \end{pmatrix} \in \boldsymbol{K}^{N_0 \times N_0}$$

とすると，$W$ はユニタリー行列になっていることが容易に確認できる．したがって，$\{W, W^*\}$ は完全再構成フィルタ・バンクである．

以下では，記述を簡単にするために $N_0 = 2^3$ の場合について，レベル 3 のサブバンド信号と多重解像度解析について述べる．この場合，用意しておくフィルタは

$$\mathbf{u}_0^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\\0\\0\\0\\0\end{pmatrix}, \quad \mathbf{u}_1^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{pmatrix},$$

$$\mathbf{u}_0^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \quad \mathbf{u}_1^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix},$$

$$\mathbf{u}_0^{(3)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad \mathbf{u}_1^{(3)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

である．$\widetilde{U}_\mu^{(j)} = U_\mu^{(j)}$ であることに注意しておく．$\mathbf{x} = (x[k])_{k=0}^{7} \in \boldsymbol{K}^8$ とする．このときサブバンド信号 $\mathbf{y}_1^{(1)}, \mathbf{y}_1^{(2)}, \mathbf{y}_1^{(3)}, \mathbf{y}_0^{(3)}$ を計算する．

$$\mathbf{y}_1^{(1)} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\mathbf{x}$$

$$= \begin{pmatrix} \frac{1}{2}x[0]\sqrt{2} - \frac{1}{2}x[1]\sqrt{2} \\ \frac{1}{2}x[2]\sqrt{2} - \frac{1}{2}x[3]\sqrt{2} \\ \frac{1}{2}x[4]\sqrt{2} - \frac{1}{2}x[5]\sqrt{2} \\ \frac{1}{2}x[6]\sqrt{2} - \frac{1}{2}x[7]\sqrt{2} \end{pmatrix}.$$

さらに

$$
\mathbf{y}_1^{(2)} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}
$$

$$
\times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{x}
$$

$$
= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \mathbf{x}
$$

$$
= \begin{pmatrix} \frac{1}{2}x[0] + \frac{1}{2}x[1] - \frac{1}{2}x[2] - \frac{1}{2}x[3] \\ \frac{1}{2}x[4] + \frac{1}{2}x[5] - \frac{1}{2}x[6] - \frac{1}{2}x[7] \end{pmatrix},
$$

$$
\mathbf{y}_1^{(3)} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}
$$

$$
\times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{x}
$$

$$
= \begin{pmatrix} \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & -\frac{1}{4}\sqrt{2} & -\frac{1}{4}\sqrt{2} & -\frac{1}{4}\sqrt{2} & -\frac{1}{4}\sqrt{2} \end{pmatrix} \mathbf{x}
$$

$$
= \left( \frac{\sqrt{2}}{4} \left( x[0] + x[1] + x[2] + x[3] - x[4] - x[5] - x[6] - x[7] \right) \right),
$$

$$\mathbf{y}_1^{(3)} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \mathbf{x}$$

$$= \begin{pmatrix} \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} \end{pmatrix} \mathbf{x}$$

$$= \left( \frac{\sqrt{2}}{4} \left( x[0] + x[1] + x[2] + x[3] + x[4] + x[5] + x[6] + x[7] \right) \right).$$

次に多重解像度分解

$$I_8 = P_0^{(3)} + P_1^{(3)} + P_1^{(2)} + P_1^{(1)}$$

を与える射影行列を求める.

$$P_1^{(1)} = U_1^{(1)} U_1^{(1)*}$$

$$= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

である. また

$$P_1^{(2)} = U_0^{(1)} U_1^{(2)} U_1^{(2)*} U_0^{(1)*}$$

$$= \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

$$P_1^{(3)} = U_0^{(1)} U_0^{(2)} U_1^{(3)} U_1^{(3)*} U_0^{(2)*} U_0^{(1)*}$$

$$= \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix},$$

$$P_0^{(3)} = U_0^{(1)} U_0^{(2)} U_0^{(3)} U_0^{(3)*} U_0^{(2)*} U_0^{(1)*}$$

$$= \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

**問題 19.4** $\mathbf{x}^T = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix}$ とする．$P_1^{(1)}\mathbf{x}$, $P_1^{(2)}\mathbf{x}$, $P_1^{(3)}\mathbf{x}$, $P_0^{(3)}\mathbf{x}$ を求め，そのグラフを描け．

### 19.5.2 ドブシー・ウェーブレット

ドブシーの考案したウェーブレットのうち，ドブシー 2 と呼ばれるウェーブレットを述べる．$N = 2^p$ であり，$N \geq 4$ を仮定する．このとき，

$$\mathbf{u}_0 = \frac{1}{4\sqrt{2}} \tau_{-1} \begin{pmatrix} 1+\sqrt{3} \\ 3+\sqrt{3} \\ 3-\sqrt{3} \\ 1-\sqrt{3} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{u}_1 = \frac{1}{4\sqrt{2}} \tau_{-1} \begin{pmatrix} 1-\sqrt{3} \\ -3+\sqrt{3} \\ 3+\sqrt{3} \\ -1-\sqrt{3} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

とする．このとき，$\mathbf{u}_0$, $\mathbf{u}_1$ はそれぞれスケーリング・フィルタ，ウェーブレット・フィルタになっている．さらに $\{\tau_{2k}\mathbf{u}_0\}_{k=0}^{M-1}$ は正規直交系になっている．なぜならば

$$\left(1+\sqrt{3}\right)\left(3-\sqrt{3}\right) + \left(3+\sqrt{3}\right)\left(1-\sqrt{3}\right) = 0$$

である．また $\{\tau_{2k}\mathbf{u}_1\}_{k=0}^{M-1}$ も正規直交系になっている．実際

$$\left(1-\sqrt{3}\right)\left(3+\sqrt{3}\right)+\left(-3+\sqrt{3}\right)\left(-1-\sqrt{3}\right)=0$$

である．また $\tau_{2k}\mathbf{u}_0$ は $\tau_{2k'}\mathbf{u}_1$ と直交している．なぜならば

$$\begin{aligned}
&\left(1+\sqrt{3}\right)\left(1-\sqrt{3}\right)+\left(3+\sqrt{3}\right)\left(-3+\sqrt{3}\right)+\left(3-\sqrt{3}\right)\left(3+\sqrt{3}\right)\\
&\qquad+\left(1-\sqrt{3}\right)\left(-1-\sqrt{3}\right)=0,\\
&\left(3-\sqrt{3}\right)\left(1-\sqrt{3}\right)+\left(1-\sqrt{3}\right)\left(-3+\sqrt{3}\right)=0,\\
&\left(1+\sqrt{3}\right)\left(3+\sqrt{3}\right)+\left(3+\sqrt{3}\right)\left(-1-\sqrt{3}\right)=0,\\
&\left(1+\sqrt{3}\right)\left(3+\sqrt{3}\right)+\left(3+\sqrt{3}\right)\left(-1-\sqrt{3}\right)+\left(3-\sqrt{3}\right)\left(1-\sqrt{3}\right)\\
&\qquad+\left(1-\sqrt{3}\right)\left(-3+\sqrt{3}\right)=0
\end{aligned}$$

となっているからである．ゆえに

$$W=\begin{pmatrix}U_0^*\\U_1^*\end{pmatrix}$$

とすると，$\{W,W^*\}$ は完全再構成な 2 分割最大間引きフィルタ・バンクをなしている．

**問題 19.5**　$\{\tau_1\mathbf{u}_0,\tau_1\mathbf{u}_1\}$ の $\boldsymbol{K}^8$ に対するレベル 2 の多重解像度分解

$$I_8=P_0^{(2)}+P_1^{(2)}+P_1^{(1)}$$

を与える射影行列を求めよ．

### 19.5.3　3 チャネル 2 間引きフィルタ・バンクの例

3 チャネル 2 間引き完全再構成フィルタ・バンクの例を一つあげておく．

$$u_0[k]=\begin{cases}\dfrac{\sqrt{2}}{4} & k=N-1,1\\[2mm] \dfrac{\sqrt{2}}{2} & k=0\\[2mm] 0 & \text{その他}\end{cases}$$

$$u_1[k]=\begin{cases}\dfrac{1}{2} & k=N-1\\[2mm] -\dfrac{1}{2} & k=1\\[2mm] 0 & \text{その他}\end{cases},\quad u_2[k]=\begin{cases}\dfrac{\sqrt{2}}{4} & k=N-1,1\\[2mm] -\dfrac{\sqrt{2}}{2} & k=0\\[2mm] 0 & \text{その他}\end{cases}$$

とする．このとき $\{\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2; \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2\}$ は 2 間引き完全再構成フィルタ・バンクをなしていることが容易に証明できる．

ウェーブレットの例はこのほかにもいろいろある．興味のある読者は [新井 2] などを参照してほしい．

## 19.6 ウェーブレットの応用例 (特異性の検出)

ウェーブレットの応用例は 1980 年代以降さまざまな分野で数多く得られている．たとえば画像圧縮，ノイズ除去などではウェーブレットの極めて有効な応用がなされている．ここでは，信号の特異性の検出への応用例を挙げておく．図 19.5 の最上段は正弦波 (sin) の一部分のグラフである．これを前節で述べたドブシー 2 でレベル 3 の多重解像度分解を行う．結果は図 19.5 のようになる．ここで A3 はレベル 3 の近似部分，D$j$ はレベル $j$ の詳細部分を表示している ($j = 1, 2, 3$).

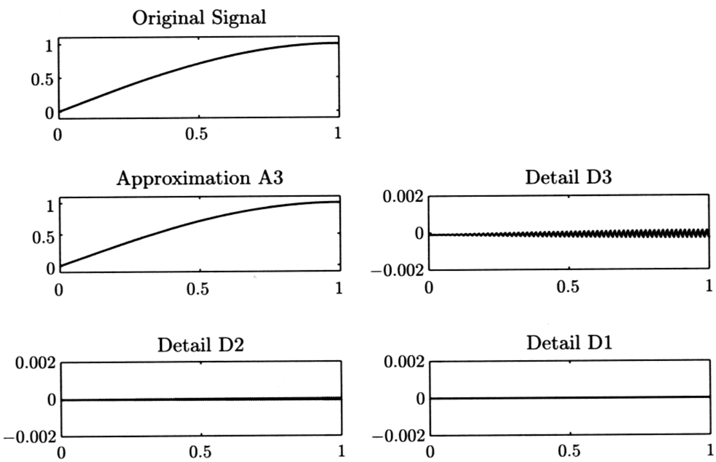



**図 19.5** 正弦波の一部の多重解像度解析

次に，この正弦波の 0.5 の部分に図 19.6 の左上図のように段差をつける．図 19.6 の右上図が段差をつけた部分の拡大図である．これをドブシー 2 で多重解像度分解すると，図 19.6 の第 2 段，第 3 段の図のようになる．この図からウェーブレットによってどこに特異点があるかが解析できることがわかるであろう．

ウェーブレットの各種の応用について本書ではこれ以上触れないが，さらにこ

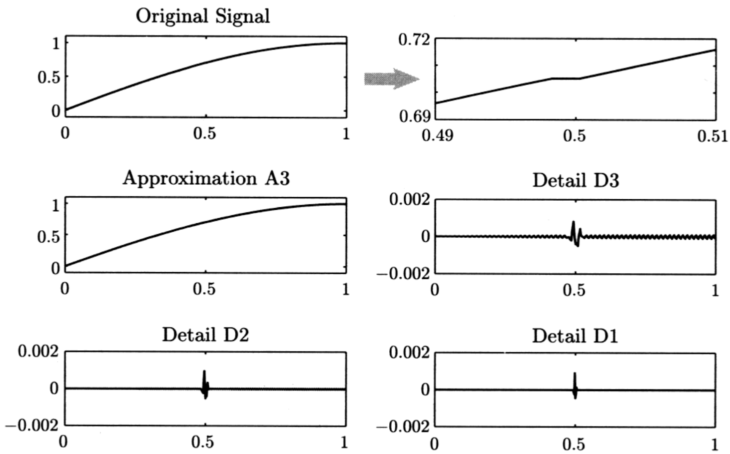



図 19.6

の方面を学習したい読者にはたとえばウェーブレットの応用についての総合報告として [たのしみ] がある．またウェーブレットの詳細，特に本章に続く有限データのウェーブレット解析とその応用については [新井 2] を参照してほしい．

# 第20章　整数値行列とその応用

この章では整数を成分とする行列について学ぶ．このような行列を**整数値行列**という．特にスミス標準形について学ぶ．また整数値行列の格子の生成への応用について述べる．

**ポイント 20**　本章は整数値の線形代数への入門である．整数の枠組みの中で線形代数を研究することは，さらに可換環を係数とする線形代数のモデルケースにもなっている．

## 20.1　スミス標準形

$\boldsymbol{Z}$ は整数全体のなす集合である．また自然数 $M, N$ に対して，

$$\boldsymbol{Z}^M = \left\{\mathbf{x} : \mathbf{x} = (x_j) \in \boldsymbol{K}^M,\ x_j \in \boldsymbol{Z}, \quad j = 1, \cdots, M\right\}$$

$$\boldsymbol{Z}^{M\times N} = \left\{A : \begin{array}{c} A = (a_{jk}) \in \boldsymbol{K}^{M\times N}, \quad a_{jk} \in \boldsymbol{Z}, \\ j = 1, \cdots, M, \quad k = 1, \cdots, N \end{array}\right\}$$

とする．11.3 節では，階数 $r$ の $(M, N)$ 行列 $A$ に対して，ある $M$ 次正則行列 $P$ と $N$ 次正則行列 $Q$ が存在し，

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

とできることを証明した (系 11.18)．ここでは整数値行列に対する標準化であるスミス標準化定理を証明する．定理を述べるために，整数値ベクトル，整数値行列に関するいくつかの基本概念を定義する．

**定義 20.1**　$\mathbf{x}_1, \cdots, \mathbf{x}_n \in \boldsymbol{Z}^N$ と $a_1, \cdots, a_n \in \boldsymbol{Z}$ に対して，$\sum_{j=1}^{n} a_j \mathbf{x}_j$ を $\mathbf{x}_1, \cdots,$

$\mathbf{x}_n$ の **$\boldsymbol{Z}$-線形結合**という．$\mathbf{x}_1, \cdots, \mathbf{x}_n$ の $\boldsymbol{Z}$-線形結合 $\sum_{j=1}^{n} a_j \mathbf{x}_j$ が $\mathbf{0}$ となるのは，$a_1 = \cdots = a_n = 0$ の場合に限るとき $\mathbf{x}_1, \cdots, \mathbf{x}_n$ は **$\boldsymbol{Z}$-線形独立**であるといい，そうでないとき **$\boldsymbol{Z}$-線形従属**という．

**定義 20.2**　$A \in \boldsymbol{Z}^{M\times M}$ が $\boldsymbol{Z}$-正則であるとは，ある $B \in \boldsymbol{Z}^{M\times M}$ で，$AB = I_M$ となるものが存在することである．

次のことが成り立つ．

**定理 20.1**　$A \in \boldsymbol{Z}^{M\times M}$ とする．$A$ が $\boldsymbol{Z}$-正則であるための必要十分条件は $|\det A| = 1$ となることである．

**証明.** $A$ が $\boldsymbol{Z}$-正則ならば，ある $B \in \boldsymbol{Z}^{M\times M}$ が存在して，$AB = I_M$ なので，

$$(\det A)(\det B) = 1$$

で，$\det A, \det B \in \boldsymbol{Z}$ であるから，$|\det A| = |\det B| = 1$ でなければならない．逆に $|\det A| = 1$ ならば，$A$ の逆行列が存在し，しかも定理 6.3 と $|\det A| = 1$ と $A$ の余因子が整数であることより，逆行列の各成分は整数になる．■

次に階数に相当する概念を定義する．

**定義 20.3**　$A \in \boldsymbol{Z}^{M\times N}$ の $r$ 次小行列式の中に 0 でないものが存在するような最大の自然数 $r$ を $A$ の**行列式階数**といい，$\mathrm{rank}_{\boldsymbol{Z}}(A)$ と表す．

次のことが成り立つ．

**定理 20.2**　$A, B \in \boldsymbol{Z}^{M\times N}$ とする．もしある二つの $\boldsymbol{Z}$-正則行列 $P \in \boldsymbol{Z}^{M\times M}$，$Q \in \boldsymbol{Z}^{N\times N}$ が存在し，$PAQ = B$ であるとする．このとき，

$$\mathrm{rank}_{\boldsymbol{Z}}(A) = \mathrm{rank}_{\boldsymbol{Z}}(B).$$

**証明.** 定理 7.5 より，$PA$ の $s$ 次小行列式は $A$ の $s$ 次小行列式の $\boldsymbol{Z}$-線形結合で表されるから，$r = \mathrm{rank}_{\boldsymbol{Z}}(A)$ とすると，$PA$ の $r+1$ 次のすべての小行列式は 0 でなければならない．したがって，$\mathrm{rank}_{\boldsymbol{Z}}(A) \geq \mathrm{rank}_{\boldsymbol{Z}}(PA)$ である．また $PAQ$ の $r$ 次小行列式は $PA$ の $r$ 次小行列式の $\boldsymbol{Z}$-線形結合となり，

$$\begin{aligned} \mathrm{rank}_{\boldsymbol{Z}}(A) \geq \mathrm{rank}_{\boldsymbol{Z}}(PA) &\geq \mathrm{rank}_{\boldsymbol{Z}}(PAQ) \\ &= \mathrm{rank}_{\boldsymbol{Z}}(B) \end{aligned}$$

である．$A = P^{-1}BQ^{-1}$ であるから，同様の考察により，$\mathrm{rank}_{\boldsymbol{Z}}(B) \geq \mathrm{rank}_{\boldsymbol{Z}}(A)$ も得る．■

スミスの標準化定理は次のものである．なおここで整数 $a, b$ に対して，$b = qa$ をみたす整数 $q$ が存在するとき，$a$ は $b$ を割り切るといい，$a|b$ と表す．

**定理 20.3（スミスの標準化）** $A \in \boldsymbol{Z}^{M\times N}$ とし，$r = \mathrm{rank}_{\boldsymbol{Z}}(A) \geq 1$ とする．このとき，ある二つの $\boldsymbol{Z}$-正則行列 $P \in \boldsymbol{Z}^{M\times M}$, $Q \in \boldsymbol{Z}^{N\times N}$ が存在し，次をみたす．

$$PAQ = \mathrm{diag}_{M,N}(\alpha_1, \cdots, \alpha_L), \quad L = \min\{M, N\},$$
$$\alpha_1, \cdots, \alpha_L \in \{0, 1, 2, \cdots\},$$
$$\alpha_1|\alpha_2, \ \ \alpha_2|\alpha_3, \ \cdots, \ \alpha_{L-1}|\alpha_L,$$
$$\alpha_1 > 0, \cdots, \alpha_r > 0, \quad \alpha_{r+1} = \cdots = \alpha_L = 0.$$

さらに $\alpha_1, \cdots, \alpha_L$ は $A$ に対して一意的に定まる．

**定義 20.4** この定理 20.3 の $\mathrm{diag}_{M,N}(\alpha_1, \cdots, \alpha_L)$ を $A$ の**スミス標準形**という．

以下では，定理 20.3 の証明を行う．ここではそのためにいくつかの準備をしておく．

整数値行列の基本変形としてつぎの操作を考える．

- I　二つの行を入れ替える．
- II　ある行を $-1$ 倍する．
- III　ある行を整数倍して別の行に加える．
- I′　二つの列を入れ替える．
- II′　ある列を $-1$ 倍する．
- III′　ある列を整数倍して別の列に加える．

これらは行列の行に関する基本変形と列に関する基本変形の特別な場合で，(4.15) より，操作 I が $F_{ij}$，操作 II が $E_i(-1)$，操作 III が $E_{i,j}(c)$ ($c$ は整数) を変形したい行列の左側から掛けることにより求める変形が得られる．また列の場合は右から掛ければよい．さらに，これらの行列は $\boldsymbol{Z}$-正則行列である．

整数値行列 $A = (a_{mn}) \in \boldsymbol{Z}^{M\times N}$ に対して次の変形を考える．$1 \leq k \leq M-1$ とする．以下に定義される操作を $\boxed{C_k}$ で表す．

$1_k$　$a_{k,k} = a_{k+1,k} = \cdots = a_{M,k} = 0$ ならば操作 $\boxed{C_k}$ を終了する．そうでなければ $2_k$ へ進む．

$2_k$　必要ならば操作 II を施して，第 $(k,k)$ 成分から第 $(M,k)$ 成分がすべて非負になるようにし，それを改めて $A = (a_{mn})$ とおく．

$3_k$　$a_{k,k}, \cdots, a_{M,k}$ の中で 0 でない最小の成分が $(k,k)$ 成分の位置にくるように必要ならば操作 I を行う．入れ替えた行列を改めて $A = (a_{mn})$ とおき，操作 $4_k$ に進む．

$4_k$　$k+1 \leq j \leq M$ に対して，$a_{j,k} = qa_{k,k} + r$ $(q, r \in \boldsymbol{Z},\ 0 \leq r < a_{k,k})$ であれば，第 $k$ 行の $-q$ 倍を第 $j$ 行に加える ( 操作 III )．この結果得られる行列を改めて $A = (a_{mn})$ とおき，操作 $5_k$ に進む．

$5_k$　$a_{k+1,k} = \cdots = a_{M,k} = 0$ ならば操作 $\boxed{C_k}$ は終了する．そうでない場合，$a_{k,k}$ が $a_{k+1,k}, \cdots, a_{M,k}$ のうち 0 でないものの最大公約数ならば操作 $6_k$ に進み，最大公約数でないならば操作 $3_k$ に戻る．

$6_k$　第 $k$ 行の整数倍を第 $j$ 行 $(k+1 \leq j \leq M)$ に加えて，第 $(j,k)$ 成分が 0 となるようにする．操作 $\boxed{C_k}$ を終了する．

整数値行列 $A = (a_{mn}) \in \boldsymbol{Z}^{M \times N}$ に対して次の変形を考える．$1 \leq k \leq N-1$ とする．以下に定義される操作を $\boxed{R_k}$ で表す．

$1'_k$　$a_{k,k} = a_{k,k+1} = \cdots = a_{k,N} = 0$ ならば操作 $\boxed{R_k}$ は終了する．そうでなければ操作 $2'_k$ に進む．

$2'_k$　必要ならば操作 II$'$ を施して，第 $(k,k)$ 成分から第 $(k,N)$ 成分がすべて非負になるようにし，それを改めて $A = (a_{mn})$ とおく．

$3'_k$　$a_{k,k}, \cdots, a_{k,N}$ の中で 0 でない最小の成分が $(k,k)$ 成分の位置にくるように必要ならば操作 I$'$ を行う．入れ替えた成分を改めて $A = (a_{mn})$ とおき，操作 $4'_k$ に進む．

$4'_k$　$k+1 \leq j \leq N$ に対して，$a_{k,j} = qa_{k,k} + r\,(q, r \in \boldsymbol{Z},\ 0 \leq r < a_{k,k})$ であれば，第 $k$ 列の $-q$ 倍を第 $j$ 列に加える ( 操作 III$'$)．この結果得られる行列を改めて $A = (a_{mn})$ とおき，操作 $5'_k$ に進む．

$5'_k$　$a_{k,k+1} = \cdots = a_{k,N} = 0$ ならば操作 $\boxed{R_k}$ を終了する．そうでない場合，$a_{k,k}$ が $a_{k,k+1}, \cdots, a_{k,N}$ の最大公約数ならば操作 $6'_k$ に進み，最大公約数でないならば操作 $3'_k$ に戻る．

$6'_k$　第 $k$ 列の整数倍を第 $j$ 列 $(k+1 \leq j \leq M)$ に加えて，第 $j$ 成分が 0 とな

るようにする．操作 $\boxed{R_k}$ を終了する．

$k=1$ とする．

1.　$k \geq M$ ならば操作 3 に進む．$k < M$ ならば操作 $\boxed{C_k}$ を行う．これにより得られる行列を改めて $A=(a_{mn})$ とおき，操作 2 に進む．

2.　$k \geq N$ ならば操作を停止する．$k < N$ ならば操作 $\boxed{R_k}$ を行う．これにより得られた行列を改めて $A=(a_{mn})$ とおき，操作 3 に進む．

3.　$a_{k+1,1} = \cdots = a_{M,1} = a_{1,k+1} = \cdots = a_{1,N} = 0$ ならば次の操作 4 に進む．そうでないならば操作 1 に戻る (この段階で $a_{k,k}$ は減少しているから，有限回の操作で必ず操作 4 に進むことができる)．

4.　$a_{k,k}|a_{r,s}\,(r=k+1,\cdots,M,\ s=k+1,\cdots,N)$ ならば操作 5 に進む．ある $a_{r,s}$ に対して，$a_{k,k} \nmid a_{r,s}$ ならば，第 $r$ 行を第 1 行に加えたものを $A=(a_{mn})$ と改めておき，操作 1 に戻る (この場合，操作 1 から操作 4 に再び来たときに得られる $a_{k,k}$ はもとの $a_{k,k}$ より小さくなっている．したがって有限回のループで次の操作 5 に移ることができる)．

5.　$k+1$ を改めて $k$ とおき，操作 1 に進む．

以上の操作により，行列 $A$ は

$$\mathrm{diag}_{M,N}\,(a_{11},\cdots,a_{LL}) \qquad (ただし，L=\min\{M,N\})$$

の形に変形される．さらにこの操作の仕方から，

$$a_{11}|a_{22},\ a_{22}|a_{33},\ \cdots,\ a_{L-1,L-1}|a_{L,L}$$

が得られる．

以上の操作は操作 I, II, III, I′, II′, III′ のみを使って実行するものであり，したがってある $\boldsymbol{Z}$-正則行列を $A$ の右と左から掛けることによって可能である．

これで定理 20.3 の前半の証明が終了した．

スミス標準化の一意性を証明する．そのための準備から始める．$A$ の $r$ 次小行列式の最大公約数を $A$ の $r$ **次行列式因子**といい，$d_r(A)$ で表す．次のことが成り立つ．

**補題 20.4**　$A \in \boldsymbol{Z}^{M\times N}$ とし，$r=\mathrm{rank}_{\boldsymbol{Z}}(A)$ とする．$1 \leq k \leq r$ のとき，$d_k(A) \neq 0$ である．また $1 \leq k < r$ のとき，$d_k(A)|d_{k+1}(A)$ である．

**証明．** もし $d_k(A)=0$ ならば $A$ のすべての $k$ 次小行列式は 0 となり，$k \leq r$ に反

する．$k+1$ 次小行列式は $k$ 次小行列式の $\boldsymbol{Z}$-線形結合で表せるので，$d_k(A)|d_{k+1}(A)$ である．■

**補題 20.5**　$A, B \in \boldsymbol{Z}^{M\times N}$ とする．もしある二つの $\boldsymbol{Z}$-正則行列 $P \in \boldsymbol{Z}^{M\times M}$, $Q \in \boldsymbol{Z}^{N\times N}$ が存在し，$PAQ = B$ であるとする．このとき，$1 \le k \le \mathrm{rank}_{\boldsymbol{Z}}(A)$ $(= \mathrm{rank}_{\boldsymbol{Z}}(B))$ をみたす $k$ に対して

$$d_k(A) = d_k(B).$$

**証明.** 定理 20.2 の証明より $B = PAQ$ の $k$ 次小行列は $A$ の $r$ 次小行列式の $\boldsymbol{Z}$-線形結合として表せる．したがって，$d_k(A)|d_k(B)$ である．また $A = P^{-1}BQ^{-1}$ に対して $d_k(B)|d_k(A)$ である．ゆえに $d_k(A) = d_k(B)$ が得られる．■

以上の補題を用いてスミス標準形の一意性を証明する．$B = PAQ$ を $A$ のスミス標準化とする．$d_k(B) = \alpha_1 \cdots \alpha_k$ $(k \le r)$ である．補題 20.5 より $d_k(A) = \alpha_1 \cdots \alpha_k$ である．これより $\alpha_1 = d_1(A)$ が得られる．これより $\alpha_2 = d_2(A)/d_1(A)$, $\cdots$, $\alpha_r = d_r(A)/d_{r-1}(A)$ となる．$r = \mathrm{rank}_{\boldsymbol{Z}}(A)$ であるから，$\alpha_1, \cdots, \alpha_L$ は $A$ にのみ依存して決まっていることがわかる (定理 20.3 の証明終わり)．

いま整数行列の行に対する基本操作 $E_1, \cdots, E_k$, 列に対する基本操作 $F_1, \cdots, F_l$ によって行列 $A \in \boldsymbol{Z}^{M\times N}$ がスミス標準形 $D$ に変形されたとする．このとき，

$$D = E_k \cdots E_1 A F_1 \cdots F_l = E_k \cdots E_1 I_M A I_N F_1 \cdots F_l$$

であるから，

$$P = E_k \cdots E_1 I_M, \quad Q = I_N F_1 \cdots F_l$$

により $PAQ = D$ を得る．この考察をもとに，次のようにして具体的なスミス標準形を求めることができる．

**例題 20.1**　$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ のスミス標準形を求めよ．

**答.** 逆行列を求めたときと同様の記法を用いる (例 4.3 参照)．

$$\left(\begin{array}{cc|cc|cc} 1 & 0 & 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc|cc} 0 & 1 & 1 & 2 & 1 & 0 \\ 1 & 0 & 2 & 1 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|cc|cc} 0 & 1 & 1 & 2 & 1 & 0 \\ 1 & -2 & 0 & -3 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc|cc} 0 & 1 & 1 & 2 & 1 & 0 \\ -1 & 2 & 0 & 3 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|cc|cc} 0 & 1 & 1 & 0 & 1 & -2 \\ -1 & -2 & 0 & 3 & 0 & 1 \end{array}\right).$$

以上より

$$\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

■

**例題 20.2** $\begin{pmatrix} 1 & 3 & 0 \\ 2 & 4 & 1 \end{pmatrix}$ のスミス標準形を求めよ．

**答．**

$$\left(\begin{array}{cc|ccc|ccc} & & & & & 1 & 0 & 0 \\ 1 & 0 & 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 2 & 4 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|ccc|ccc} & & & & & 1 & 0 & 0 \\ 1 & 0 & 1 & 3 & 0 & 0 & 1 & 0 \\ -2 & 1 & 0 & -2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|ccc|ccc} & & & & & 1 & -3 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -2 & 1 & 0 & -2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|ccc|ccc} & & & & & 1 & -3 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 & 2 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|ccc|ccc} & & & & & 1 & 0 & -3 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ -2 & 1 & 0 & 1 & 0 & 0 & 1 & 2 \end{array}\right).$$

したがって求めるスミス標準形は次のものである．

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

■

## 20.2　整数値行列による格子の生成

本節は整数値行列の応用として整数行列から作られる格子について述べる．正則な整数値行列 $R$，すなわち $R \in \boldsymbol{Z}^{N\times N}$ かつ $\det R \neq 0$ をみたす $R$ に対して

$$LAT(R) = \left\{R\mathbf{x} : \mathbf{x} \in \boldsymbol{Z}^N\right\}$$

とおき，$R$ によって作られる**格子**という．$LAT(R) \subset \boldsymbol{Z}^N$ である．特に $R = I_N$ のとき，$LAT(I_N) = \boldsymbol{Z}^N$ を**基本格子**という．$N = 2$ の場合に格子のいくつかの代表例をあげる．

**例 20.1（長方形格子）**　$a, b$ を正の整数とし，

$$R = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

とする．このとき $LAT(R) = \left\{ \begin{pmatrix} an \\ bm \end{pmatrix} : n, m \in \boldsymbol{Z} \right\}$ である．これは長方形格子と呼ばれている．

**例 20.2（六角格子）**

$$R = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$$

に対して，$LAT(R) = \left\{ \begin{pmatrix} m+n \\ 2\,(m-n) \end{pmatrix} : n, m \in \boldsymbol{Z} \right\}$ は六角格子と呼ばれている．

**例 20.3（五の目格子）**

$$R = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

に対して，$LAT(R) = \left\{ \begin{pmatrix} m+n \\ n-m \end{pmatrix} : n, m \in \boldsymbol{Z} \right\}$ を五の目格子という．

一般に行列 $A \in \boldsymbol{K}^{M \times N}$ と集合 $X \subset \boldsymbol{K}^N$ に対して,

$$A(X) = \{A\mathbf{x} : \mathbf{x} \in X\}$$

と表す．次の定理が成り立つ．

**定理 20.6**　$R \in \boldsymbol{Z}^{N \times N}$ を正則，また $U \in \boldsymbol{Z}^{N \times N}$ を $\boldsymbol{Z}$-正則であるとする．このとき,

$$LAT(RU) = LAT(R), \tag{20.1}$$

$$LAT(UR) = U(LAT(R)). \tag{20.2}$$

**証明.** $\mathbf{x} \in LAT(R)$ とすると，ある $\mathbf{y} \in \boldsymbol{Z}^N$ で $\mathbf{x} = R\mathbf{y}$ となるものが存在する．$\mathbf{z} = U^{-1}\mathbf{y}$ とおくと，$\mathbf{z} \in \boldsymbol{Z}^N$ であり，$\mathbf{x} = RU\mathbf{z} \in LAT(RU)$ となっている．一方，明らかに $LAT(RU) \subset LAT(R)$ である．

$LAT(UR) \ni \mathbf{x}$ とすると，$\mathbf{x} = UR\mathbf{y}$ となる $\mathbf{y} \in \boldsymbol{Z}^N$ が存在する．$\mathbf{z} = R\mathbf{y}$ とおくと, $\mathbf{z} \in LAT(R)$ であるから, $\mathbf{x} \in U(LAT(R))$ である．一方, $\mathbf{x} \in U(LAT(R))$ ならば, ある $\mathbf{y} \in \boldsymbol{Z}^N$ が存在し, $\mathbf{x} = UR\mathbf{y}$ が成り立っているので, $\mathbf{x} \in LAT(UR)$ である．（注：(20.2) は $U$ が $\mathbf{Z}$-正則でなくとも成り立つ．）■

(20.1) より，$S = \begin{pmatrix} 2 & 3 \\ 0 & -2 \end{pmatrix}$ に対して，$LAT(S)$ も六角格子であることがわかる．なぜなら

$$U = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

は $\det(U) = 1$ より $\boldsymbol{Z}$-正則であり,

$$S = \begin{pmatrix} 2 & 3 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

となっている．

定理 20.6 より次のことがわかる．

**定理 20.7**　$R \in \boldsymbol{Z}^{N \times N}$ を正則とし，$PRQ = D$ をスミス標準形とする．このとき,

$$LAT(R) = P^{-1}(LAT(D)).$$

**証明.** $LAT(PRQ) = LAT(PR) = P(LAT(R))$ である．■

**例 20.4**　$R = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$ とする．$P = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ とすると,

$$PRQ = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

である．$P^{-1} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$ であるから，六角格子は

$$\begin{aligned} LAT(R) &= \left\{ \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} m \\ 4n \end{pmatrix} : m, n \in \boldsymbol{Z} \right\} \\ &= \left\{ \begin{pmatrix} m \\ 2m - 4n \end{pmatrix} : m, n \in \boldsymbol{Z} \right\} \end{aligned}$$

となることがわかる．

**定義 20.5**　正則行列 $R_1$, $R_2 \in \boldsymbol{Z}^{N \times N}$ が

$$LAT(R_2) \subset LAT(R_1)$$

をみたすとき，$R_2$ は $LAT(R_1)$ の**部分格子を生成する**という．また $LAT(R_2)$ を $LAT(R_1)$ の**部分格子**という．

次のことが成り立つ．

**定理 20.8**　$R_1$, $R_2 \in \boldsymbol{Z}^{N \times N}$ を正則行列とする．次の (i), (ii) は同値である．

(i)　$LAT(R_2) \subset LAT(R_1)$.

(ii)　$R_1^{-1} R_2 \in \boldsymbol{Z}^{N \times N}$.

**証明.** (i) ⇒ (ii)　任意の $\mathbf{n} \in \boldsymbol{Z}^N$ に対して，$R_2\mathbf{n} = R_1\mathbf{m}$ をみたす $\mathbf{m} \in \boldsymbol{Z}^N$ が存在する．したがって，$R_1^{-1}R_2\mathbf{n} = \mathbf{m} \in \boldsymbol{Z}^N$ であるから，$R_1^{-1}R_2$ は整数値行列である．

(ii) ⇒ (i)　任意の $\mathbf{n} \in \boldsymbol{Z}^N$ に対して，$R_1^{-1}R_2\mathbf{n} = \mathbf{m} \in \boldsymbol{Z}^N$ であるから，$R_2\mathbf{n} = R_1\mathbf{m} \in LAT(R_1)$ である．■

$LAT(R_2)$ が $LAT(R_1)$ の部分格子であるとする．$\mathbf{x}, \mathbf{y} \in LAT(R_1)$ に対して，

$$\mathbf{x} \sim \mathbf{y} \Longleftrightarrow \mathbf{x} - \mathbf{y} \in LAT(R_2)$$

と定義する．$\mathbf{x} \in LAT\,(R_1)$ に対して $[\mathbf{x}]_{R_1,R_2} = \{\mathbf{y} \in LAT(R_1) : \mathbf{x} \sim \mathbf{y}\}$ と定義し，

$$LAT(R_1)/LAT(R_2) = \left\{[\mathbf{x}]_{R_1,R_2} : \mathbf{x} \in LAT(R_1)\right\}$$

とする．ここで $\mathbf{x}, \mathbf{y} \in LAT(R_1)$ かつ $\mathbf{x} \sim \mathbf{y}$ ならば

$$[\mathbf{x}]_{R_1,R_2} = [\mathbf{y}]_{R_1,R_2}$$

であることに注意する．

特に $[\mathbf{x}]_{I_N,R_2} = [\mathbf{x}]_{R_2}$ と表す．

**注意 20.1**　$\sim$ は

(i)　$\mathbf{x} \sim \mathbf{x}$,

(ii)　$\mathbf{x} \sim \mathbf{y} \Longleftrightarrow \mathbf{y} \sim \mathbf{x}$,

(iii)　$\mathbf{x} \sim \mathbf{y}$ かつ $\mathbf{y} \sim \mathbf{z}$ ならば $\mathbf{x} \sim \mathbf{z}$

の三つの性質をみたすことが容易に証明される．このような性質をみたす関係を一般に**同値関係**という．$[\mathbf{x}]_{R_1,R_2}$ は $\mathbf{x}$ の**同値類**，$LAT(R_1)/LAT(R_2)$ は**商集合**と呼ばれているものである．

**定義 20.6**

$$[0,1)^N = \left\{x : \mathbf{x} = (x_j) \in \boldsymbol{R}^N, \quad 0 \le x_j < 1,\ j = 1, \cdots, N\right\}$$

とする．$R \in \boldsymbol{Z}^{N \times N}$ に対して，

$$FPD(R) = \left\{R\mathbf{x} : \mathbf{x} \in [0,1)^N\right\}$$

を**基本平行体**という．

基本平行体は立方体 $[0,1)^N$ の $R$ による像に他ならない．$FDP(R) \subset \boldsymbol{R}^N$ である．次のことが成り立つ．

**定理 20.9**　$R \in \boldsymbol{Z}^{N \times N}$ と $U \in \boldsymbol{Z}^{N \times N}$ に対して，

$$FPD(UR) = U\left(FPD(R)\right).$$

**証明．** $FPD(UR) \ni \mathbf{x}$ とすると，$\mathbf{x} = UR\mathbf{y}$ となる $\mathbf{y} \in [0,1)^N$ が存在する．$\mathbf{z} = R\mathbf{y}$ とおくと，$\mathbf{z} \in FPD(R)$ であるから，$\mathbf{x} \in U\ (FPD(R))$ である．一方，$\mathbf{x} \in U\ (FPD(R))$ ならば，ある $\mathbf{y} \in [0,1)^N$ が存在し，$\mathbf{x} = UR\mathbf{y}$ が成り立っているので，$\mathbf{x} \in FPD(UR)$ である．■

次にシフト・ベクトルを定義する．格子は部分格子をシフト・ベクトルでシフトして作られることを示す．

**定義 20.7**　$R_1, R_2 \in \boldsymbol{Z}^{N\times N}$ を正則行列で，$LAT(R_2) \subset LAT(R_1)$ であるとする．このとき，

$$S(R_1, R_2) = LAT(R_1) \cap FPD(R_2)$$

とし，これに属するベクトルを**シフト・ベクトル**という．特に $S(R_2) = S(I_N, R_2)$ と表す．

**定理 20.10**　$R_1, R_2 \in \boldsymbol{Z}^{N\times N}$ を正則行列で，$LAT(R_2) \subset LAT(R_1)$ とする．このとき，任意の $\mathbf{x} \in LAT(R_1)$ に対して，ある $\mathbf{a} \in S(R_1, R_2)$ とある $\mathbf{p} \in LAT(R_2)$ で，

$$\mathbf{x} = \mathbf{a} + \mathbf{p}$$

となるものが一意的に存在する．

**証明.** 証明のためいくつか記号を準備する．集合 $X, Y \subset \boldsymbol{R}^N$ に対して，

$$X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X,\ \mathbf{y} \in Y\}$$

とし，特に $\mathbf{y} \in \boldsymbol{R}^N$ に対して，$X + \mathbf{y} = X + \{\mathbf{y}\}$ と表す．

$$\bigcup_{\mathbf{n}\in\boldsymbol{Z}^N} \left([0,1)^N + \mathbf{n}\right) = \boldsymbol{R}^N$$

である．$R_2$ は正則であるから，

$$\begin{aligned}\bigcup_{\mathbf{n}\in\boldsymbol{Z}^N} R_2\left([0,1)^N + \mathbf{n}\right) &= R_2\left(\bigcup_{\mathbf{n}\in\boldsymbol{Z}^N}\left([0,1)^N + \mathbf{n}\right)\right)\\ &= R_2\left(\boldsymbol{R}^N\right) = \boldsymbol{R}^N\end{aligned}$$

である．$R_2\left([0,1)^N + \mathbf{n}\right) = FPD(R_2) + R_2\mathbf{n}$ より

$$\begin{aligned}\boldsymbol{R}^N &= \bigcup_{\mathbf{n}\in\boldsymbol{Z}^N} R_2\left([0,1)^N + \mathbf{n}\right) = \bigcup_{\mathbf{n}\in\boldsymbol{Z}^N}\left(FPD(R_2) + R_2\mathbf{n}\right)\\ &= FPD(R_2) + LAT(R_2)\end{aligned}$$

が成り立つ．これを用いて $LAT(R_1) = S(R_1, R_2) + LAT(R_2)$ を示す．

$\mathbf{x} \in LAT(R_1)$ とするとすでに示したことから，ある $\mathbf{a} \in FDP(R_2)$ とある $\mathbf{p} \in LAT(R_2)$ が存在し，$\mathbf{x} = \mathbf{a} + \mathbf{p}$ が成り立つ．$\mathbf{a} = \mathbf{x} - \mathbf{p} \in LAT(R_1)$ であるから，$\mathbf{a} \in LAT(R_1) \cap FPD(R_2) = S(R_1, R_2)$ となっている．ゆえに

$LAT(R_1) \subset S(R_1, R_2) + LAT(R_2)$ が成り立つ．逆向きの包含関係は明らかに成り立つ．

次に一意性を示す．$\mathbf{a} + \mathbf{p} = \mathbf{a}' + \mathbf{p}'$ $(\mathbf{a}, \mathbf{a}' \in S(R_1, R_2),\ \mathbf{p}, \mathbf{p}' \in LAT(R_2))$ であるとする．$\mathbf{a} = R_2\mathbf{c}$, $\mathbf{a}' = R_2\mathbf{c}'$ $(\mathbf{c}, \mathbf{c}' \in [0,1)^N)$ と表せている．したがって，$R_2(\mathbf{c} - \mathbf{c}') = \mathbf{a} - \mathbf{a}' \in LAT(R_2)$ である．$R_2$ の正則性から，$\mathbf{c} - \mathbf{c}' \in \boldsymbol{Z}^N$ でもある．ゆえに $\mathbf{c} - \mathbf{c}' = \mathbf{0}$ でなければならない． ■

**系 20.11** $R_1, R_2 \in \boldsymbol{Z}^{N\times N}$ を正則行列で，$LAT(R_2) \subset LAT(R_1)$ とする．$\mathcal{S}(R_1, R_2) = \left\{[\mathbf{a}]_{R_1,R_2} : \mathbf{a} \in S(R_1, R_2)\right\}$ とする．このとき

$$LAT(R_1)/LAT(R_2) = \mathcal{S}(R_1, R_2).$$

**証明.** $LAT(R_1)/LAT(R_2) \subset \mathcal{S}(R_1, R_2)$ を示せば十分である．任意に $[\mathbf{x}]_{R_1,R_2} \in LAT(R_1)/LAT(R_2)$ をとる．このとき，定理 20.10 より，ある $\mathbf{a} \in S(R_1, R_2)$ とある $\mathbf{p} \in LAT(R_2)$ が存在し，$\mathbf{x} = \mathbf{a} + \mathbf{p}$ が成り立つ．したがって，$[\mathbf{x}]_{R_1,R_2} = [\mathbf{a}]_{R_1,R_2} \in \mathcal{S}(R_1, R_2)$ である． ■

特に $R_1 = I_N$ の場合は次のようになる．

**系 20.12** $R \in \boldsymbol{Z}^{N\times N}$ を正則行列とする．このとき，任意の $\mathbf{x} \in \boldsymbol{Z}^{N\times N}$ に対して，ある $\mathbf{a} \in S(R)$ とある $\mathbf{p} \in LAT(R)$ で，

$$\mathbf{x} = \mathbf{a} + \mathbf{p}$$

となるものが一意的に存在する．

**系 20.13** $R \in \boldsymbol{Z}^{N\times N}$ を正則行列とする．$\mathcal{S}(R) = \{[\mathbf{a}] : \mathbf{a} \in S(R)\}$ とすると，

$$\boldsymbol{Z}^N/LAT(R) = \mathcal{S}(R).$$

# 第 4 部
# 線形代数の抽象化

# 第21章　線形空間

これまで $\boldsymbol{K}^N$ ならびにその部分空間を扱ってきた．この章では，$\boldsymbol{K}^N$ あるいはその部分空間を抽象化した概念である線形空間について述べる．線形空間はベクトル空間ともいう．

**ポイント21**　これまでは具体的な数ベクトル空間を扱っていたが，それを抽象化した線形空間について学ぶ．抽象化することにより，線形代数の適用範囲はさらに広がる．この章では，抽象的な数学の考え方に慣れるようにしてほしい．

## 21.1　線形空間の定義と例

はじめに線形空間の定義をする．

**定義 21.1**　空でない集合 $X$ が $\boldsymbol{K}$ 上の**線形空間**あるいは**ベクトル空間**であるとは，$X$ の任意の元 $\mathbf{x}, \mathbf{y}$ に対して**和**と呼ばれる演算 $\mathbf{x}+\mathbf{y}$ が定義され，また $X$ の任意の元 $\mathbf{x}$ と $\alpha \in \boldsymbol{K}$ に対して**スカラー積**と呼ばれる演算 $\alpha\mathbf{x}$ が定義され，次の (1) ～ (7) の条件をみたすことである．

(1) (和の結合則)　$(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z}) \quad (\mathbf{x}, \mathbf{y}, \mathbf{z} \in X)$.

(2) (和の交換則)　$\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{z} \quad (\mathbf{x}, \mathbf{y} \in X)$.

(3) (零元の存在)　ある元 $\mathbf{0} \in X$ が存在し，すべての $\mathbf{x} \in X$ に対して，$\mathbf{x}+\mathbf{0} = \mathbf{x}$ となる．

(4) (逆元の存在)　任意の $\mathbf{x} \in X$ に対して，ある $\mathbf{x}' \in X$ が存在し，$\mathbf{x}+\mathbf{x}' = \mathbf{0}$ となる．

(5) (分配則)　$\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}, \quad (\alpha+\beta)\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}$

$(\mathbf{x}, \mathbf{y} \in X,\ \alpha, \beta \in \boldsymbol{K})$.

(6) (スカラー積の結合則)　$(\alpha\beta)\,\mathbf{x} = \alpha\,(\beta\mathbf{x})$　$(\mathbf{x} \in X,\ \alpha, \beta \in \boldsymbol{K})$.

(7) (単位元の存在)　$1\mathbf{x} = \mathbf{x}$　$(\mathbf{x} \in X)$.

特に $\boldsymbol{K} = \boldsymbol{R}$ の場合を**実線形空間**あるいは**実ベクトル空間**，$\boldsymbol{K} = \boldsymbol{C}$ の場合を**複素線形空間**あるいは**複素ベクトル空間**という．

線形空間の零元および $\mathbf{x} \in X$ に対する逆元は一意的に定まる．実際，$\mathbf{0}, \mathbf{0}'$ を零元とすると，零元の定義と和の交換則より

$$\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}$$

が得られる．また $\mathbf{x}', \mathbf{x}''$ を $\mathbf{x}$ の逆元とすると，和の交換則と結合則より

$$\begin{aligned}\mathbf{x}'' &= \mathbf{0} + \mathbf{x}'' = (\mathbf{x} + \mathbf{x}') + \mathbf{x}'' = (\mathbf{x} + \mathbf{x}'') + \mathbf{x}' \\ &= \mathbf{0} + \mathbf{x}' = \mathbf{x}'\end{aligned}$$

も得られる．

以下では，$\mathbf{x}$ の逆元は $-\mathbf{x}$ で表すことにする．

$\mathbf{x}_1, \cdots, \mathbf{x}_N \in X$ に対して，和 $\mathbf{x}_1 + \cdots + \mathbf{x}_N$ を

$$(\cdots((\mathbf{x}_1 + \mathbf{x}_2) + \mathbf{x}_3) + \cdots) + \mathbf{x}_N$$

により定義する．結合則から括弧で括る順序を変えることができる．またこの和を

$$\sum_{j=1}^{N} \mathbf{x}_j$$

と表す．

線形空間の例としては，$X = \boldsymbol{K}^N$ があげられる．また $\boldsymbol{K}^N$ の部分空間 $X$ も線形空間の例である．このほか次のような例もある．

**例 21.1**　$\mathcal{P}_N$ を高々 $N$ 次の複素係数の多項式全体のなす集合とする．すなわち

$$\mathcal{P}_N = \left\{ p : p(z) = \sum_{n=0}^{N} a_n z^n, \quad a_0, \cdots, a_N \in \boldsymbol{C} \right\}$$

とする．$\mathcal{P}_N \ni p, q$, $\boldsymbol{C} \ni \alpha$ に対して，

$$(p + q)\,(z) = p(z) + q(z), \quad (\alpha p)\,(z) = \alpha p(z)$$

により $p + q$, $\alpha p$ を定義すると，$p + q$, $\alpha p \in \mathcal{P}_N$ であり，さらにこの演算により $\mathcal{P}_N$ は複素線形空間になっている．

**例 21. 2**

$$\boldsymbol{K}^{M\times N}=\{A: A \text{ は} \boldsymbol{K} \text{に属するスカラーを成分とする} (M,N) \text{行列}\}$$

とする．このとき，$A, B \in \boldsymbol{K}^{M\times N}$, $\alpha \in \boldsymbol{K}$ に対して，$A+B$, $\alpha A$ を行列の和とスカラー積とすると，この演算により $\boldsymbol{K}^{M\times N}$ は $\boldsymbol{K}$ 上の線形空間になる．

**例 21. 3**　$X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の線形空間とし，

$$X_1\times\cdots\times X_n=\{(\mathbf{x}_1,\cdots,\mathbf{x}_n): \mathbf{x}_j\in X_j,\quad j=1,\cdots,n\}$$

をその直積集合とする．$(\mathbf{x}_1,\cdots,\mathbf{x}_n)$, $(\mathbf{y}_1,\cdots,\mathbf{y}_n) \in X_1\times\cdots\times X_n$ および $\alpha \in \boldsymbol{K}$ に対して

$$(\mathbf{x}_1,\cdots,\mathbf{x}_n)+(\mathbf{y}_1,\cdots,\mathbf{y}_n)=(\mathbf{x}_1+\mathbf{y}_1,\cdots,\mathbf{x}_n+\mathbf{y}_n),$$
$$\alpha(\mathbf{x}_1,\cdots,\mathbf{x}_n)=(\alpha\mathbf{x}_1,\cdots,\alpha\mathbf{x}_n)$$

により $X_1\times\cdots\times X_n$ は線形空間になる．これを**直積線形空間**という．

線形空間の例として，線形写像や $n$ 重線形写像全体からなる空間もしばしば利用される．それについて述べるため，まず線形写像，$n$ 重線形写像の定義をする．

**定義 21. 2**　(1)　$X, Y$ を $\boldsymbol{K}$ 上の線形空間とする．$X$ から $Y$ への写像 $f$ が**線形写像**であるとは，$\mathbf{x}, \mathbf{y} \in X$, $\alpha \in \boldsymbol{K}$ に対して，

$$f(\mathbf{x}+\mathbf{y})=f(\mathbf{x})+f(\mathbf{y}),$$
$$f(\alpha\mathbf{x})=\alpha f(\mathbf{x})$$

をみたすことである．

(2)　$X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の線形空間とし，$X_1\times\cdots\times X_n$ をその直積線形空間とする．$W$ を $\boldsymbol{K}$ 上の線形空間とする．$X_1\times\cdots\times X_n$ から $W$ への写像 $\Phi$ が $n$ **重線形**であるとは，任意の $j\in\{1,\cdots,n\}$ と任意に固定された $\mathbf{a}_i \in X_i$ (ただし $i\neq j$) に対して，写像

$$X_j \ni \mathbf{x} \longmapsto \Phi(\mathbf{a}_1,\cdots,\mathbf{a}_{j-1},\mathbf{x},\mathbf{a}_{j+1},\cdots,\mathbf{a}_n)$$

が $X_j$ から $W$ への線形写像になっていることである．このような写像を $X_1\times\cdots\times X_n$ から $W$ への $n$ **重線形写像**という．特に 1 重線形写像は線形写像を表すものとする．

以下では，$X_1\times\cdots\times X_n$ から $W$ への $n$ 重線形写像全体のなす集合を

$$L(X_1, \cdots, X_n; W)$$

により表す．$\Phi, \Psi \in L(X_1, \cdots, X_n; W)$, $\alpha \in \boldsymbol{K}$ に対して，

$$(\Phi + \Psi)(x_1, \cdots, x_n) = \Phi(x_1, \cdots, x_n) + \Psi(x_1, \cdots, x_n),$$
$$(\alpha\Phi)(x_1, \cdots, x_n) = \alpha\Phi(x_1, \cdots, x_n)$$

$((x_1, \cdots, x_n) \in X_1 \times \cdots \times X_n)$ とすると，$\Phi + \Psi, \alpha\Phi \in L(X_1, \cdots, X_n; W)$ となっている．この演算により $L(X_1, \cdots, X_n; W)$ は $\boldsymbol{K}$ 上の線形空間になっている．

**定義 21.3** $X, X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の線形空間とする．$X$ から $\boldsymbol{K}$ への線形写像を $X$ 上の**線形形式**あるいは**線形汎関数**という．また $X_1, \cdots, X_n$ から $\boldsymbol{K}$ への $n$ 重線形写像を $n$ **重線形形式**という．特に $X$ 上の線形汎関数全体のなす集合 $L(X; \boldsymbol{K})$ を $X^*$ と表し，$X$ の**双対空間**という．

**注意 21.1** 2 重線形のことは**双線形**といい，2 重線形写像を**双線形写像**，2 重線形形式を**双線形形式**という．

**定義 21.4** $X$ を $\boldsymbol{K}$ 上の線形空間とし，$\mathbf{x}_1, \cdots, \mathbf{x}_n \in X$ とする．このとき $\boldsymbol{K}$ の元 $\alpha_1, \cdots, \alpha_n$ によって $\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$ と表される $X$ の元を $\mathbf{x}_1, \cdots, \mathbf{x}_n$ の**線形結合**という．

**定義 21.5** $X$ を $\boldsymbol{K}$ 上の線形空間とし，$Y \subset X$ とする．$Y$ が $X$ の**線形部分空間**あるいは**部分空間**であるとは，任意の $\mathbf{x}, \mathbf{y} \in Y$, $\alpha \in \boldsymbol{K}$ に対して

$$\mathbf{x} + \mathbf{y} \in Y, \quad \alpha\mathbf{x} \in Y$$

をみたすことである．

$\{\mathbf{0}\}$ は $X$ の部分空間であるが，これを**自明な部分空間**という．自明でない部分空間を**非自明な部分空間**という．$X$ の部分空間自身一つの線形空間になっている．

**例 21.4** $E \subset X$ に対して，$E$ の有限個の元の線形結合全体のなす集合を $\mathrm{span}(E)$ とする．$\mathrm{span}(E)$ は $X$ の部分空間になっている．これを $E$ により**生成される部分空間**という．

部分空間から次の線形空間を作ることができる．

**例 21.5（商線形空間）** $X$ を $\boldsymbol{K}$ 上の線形空間で，$M$ をその部分空間とする．

$\mathbf{x}, \mathbf{y} \in X$ に対して，$\mathbf{x} - \mathbf{y} \in M$ であるとき，$\mathbf{x} \sim \mathbf{y}$ と表す．このとき，

$$\mathbf{x} \sim \mathbf{x},$$
$$\mathbf{x} \sim \mathbf{y} \Longrightarrow \mathbf{y} \sim \mathbf{x},$$
$$\mathbf{x} \sim \mathbf{y},\ \mathbf{y} \sim \mathbf{z} \Longrightarrow \mathbf{x} \sim \mathbf{z}$$

をみたしている．$\mathbf{x} \in X$ に対して，

$$[\mathbf{x}] = \{\mathbf{y} \in X : \mathbf{x} \sim \mathbf{y}\}$$

と定義する．もし $\mathbf{x} \sim \mathbf{y}$ ならば $[\mathbf{x}] = [\mathbf{y}]$ となっている．

$$X/M = \{[\mathbf{x}] : \mathbf{x} \in X\}$$

と定義する．$X/M \ni [\mathbf{x}], [\mathbf{y}]$ および $\alpha \in \boldsymbol{K}$ に対して，

$$[\mathbf{x}] + [\mathbf{y}] = [\mathbf{x} + \mathbf{y}],$$
$$\alpha [\mathbf{x}] = [\alpha \mathbf{x}]$$

とすると，これにより $X/M$ は $\boldsymbol{K}$ 上の線形空間となっていることが容易に証明できる．これを $X$ の $M$ に関する**商線形空間**という．

線形空間の基底は次のように定義される．

**定義 21.6**　(1)　$\mathbf{x}_1, \cdots, \mathbf{x}_n$ の線形結合 $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_n \mathbf{x}_n$ が $\mathbf{0}$ となるのは $\alpha_1 = \cdots = \alpha_n = 0$ に限るとき，$\mathbf{x}_1, \cdots, \mathbf{x}_n$ は**線形独立**であるという．線形独立でないことを**線形従属**という．

(2)　$X$ を $\boldsymbol{K}$ 上の線形空間とする．ある有限個の線形独立な元 $\mathbf{x}_1, \cdots, \mathbf{x}_N \in X$ が，もし

$$X = \mathrm{span}\,\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$$

をみたすとき，$\mathbf{x}_1, \cdots, \mathbf{x}_N$ は $X$ の**基底**であるという．あるいは $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ は $X$ の**基底**であるという．簡単に $\{\mathbf{x}_j\}_{j=1}^N$ を $X$ の基底ということもある．

次のことが成り立つ．

**命題 21.1**　$X$ を $\boldsymbol{K}$ 上の線形空間とする．もし $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ ならびに $\{\mathbf{y}_1, \cdots, \mathbf{y}_M\}$ がともに $X$ の基底であれば，$N = M$ である．

**証明.** はじめに，$M \leq N$ を示す．各 $\mathbf{y}_j$ に対して，$\mathbf{y}_j = \sum_{k=1}^{N} c_{kj} \mathbf{x}_k$ と表すことができる．$C = (c_{kj}) \in \boldsymbol{K}^{N \times M}$ とおく．いま $\mathbf{a} = (a_j) \in \boldsymbol{K}^M$ が $C\mathbf{a} = \mathbf{0}$ をみた

していたとする．このとき，

$$\sum_{j=1}^{M} a_j \mathbf{y}_j = \sum_{j=1}^{M} a_j \sum_{k=1}^{N} c_{kj} \mathbf{x}_k = \sum_{k=1}^{N} \left( \sum_{j=1}^{M} c_{kj} a_j \right) \mathbf{x}_k = \mathbf{0}$$

である．ゆえに $\mathbf{a} = \mathbf{0}$ が得られる．したがって，

$$\dim \mathcal{R}(C) = \dim \mathcal{R}(C) + \dim \mathcal{N}(C) = M$$

である．これより，$M = \operatorname{rank}(C) \leq N$ でなければならない．一方，各 $\mathbf{x}_j$ に対して，$\mathbf{x}_j = \sum_{k=1}^{M} c'_{kj} \mathbf{y}_j$ と表せるので，既述の議論と同様にして $N \leq M$ も得られる．■

**定義 21.7**　$\boldsymbol{K}$ 上の線形空間 $X$ が基底 $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ をもつとき，$X$ は**有限次元**の線形空間であるという．また $N$ を $X$ の**次元**といい，$\dim X = N$ と表す．命題 21.1 より $N$ は $X$ の基底のとり方によらない量である．$X$ のいかなる有限個の元も $X$ の基底にならないときは，$X$ は**無限次元**の線形空間であるという．

これまであげてきた例はみな有限次元線形空間である．無限次元線形空間の例としては次のものがある．

**例 21.6（無限直和）**　$X_k$ を $\boldsymbol{K}$ 上の線形空間とする $(k = 0, 1, 2, \cdots)$．このとき $X_k$ の要素 $\mathbf{x}_k$ $(k = 0, 1, 2, \cdots)$ からなる無限列 $\{\mathbf{x}_k\}_{k=0}^{\infty}$ 全体のなす集合を $\mathcal{X}$ と表す．二つの無限列 $\{\mathbf{x}_k\}_{k=0}^{\infty}$, $\{\mathbf{y}_k\}_{k=0}^{\infty} \in \mathcal{X}$ は $\mathbf{x}_k = \mathbf{y}_k$ $(k = 0, 1, 2, \cdots)$ をみたすとき等しいといい，$\{\mathbf{x}_k\}_{k=0}^{\infty} = \{\mathbf{y}_k\}_{k=0}^{\infty}$ と表す．$\{\mathbf{x}_k\}_{k=0}^{\infty}$, $\{\mathbf{y}_k\}_{k=0}^{\infty} \in \mathcal{X}$ および $\alpha \in \boldsymbol{K}$ に対して

$$\{\mathbf{x}_k\}_{k=0}^{\infty} + \{\mathbf{y}_k\}_{k=0}^{\infty} = \{\mathbf{x}_k + \mathbf{y}_k\}_{k=0}^{\infty},$$
$$\alpha \{\mathbf{x}_k\}_{k=0}^{\infty} = \{\alpha \mathbf{x}_k\}_{k=0}^{\infty}$$

とすると $\mathcal{X}$ は線形空間になる．特に

$$\oplus_{k=0}^{\infty} X_k = \{\{\mathbf{x}_k\}_{k=0}^{\infty} \in \mathcal{X} : \mathbf{x}_k \neq \mathbf{0} \text{ なる } k \text{ は有限個}\}$$

とし，$X_1, X_2, \cdots$ の**無限直和**という．$\oplus_{k=0}^{\infty} X_k$ は $\mathcal{X}$ の部分空間である．$j = 0, 1, 2, \cdots$ に対して

$$X_j^0 = \{\{\mathbf{x}_k\}_{k=0}^{\infty} \in \mathcal{X} : \mathbf{x}_k = \mathbf{0},\ k \neq j\}$$

とすると，$X_j^0$ は $\oplus_{k=0}^{\infty} X_k$ の部分空間である．また $N = 1, 2, \cdots$ に対して，

$$X_0 \oplus \cdots \oplus X_N = \{\{\mathbf{x}_k\}_{k=0}^{\infty} \in \mathcal{X} : \mathbf{x}_k = \mathbf{0},\ k > N\}$$

と表すと，$X_0 \oplus \cdots \oplus X_N$ は $\oplus_{k=0}^{\infty} X_k$ の部分空間である．

$\mathbf{x} = \{\mathbf{x}_k\}_{k=0}^{\infty} \in \oplus_{k=0}^{\infty} X_k$ を形式的に

$$\mathbf{x} = \sum_{k=0}^{\infty} \mathbf{x}_k = \mathbf{x}_1 + \mathbf{x}_2 + \cdots$$

と表す．また $\mathbf{x} = \{\mathbf{x}_k\}_{k=0}^{\infty} \in X_0 \oplus \cdots \oplus X_N$ は

$$\mathbf{x} = \sum_{k=0}^{N} \mathbf{x}_k = \mathbf{x}_1 + \cdots + \mathbf{x}_N$$

と表す．

**問題 21.1**　$X_j$ を $\{\mathbf{0}\}$ ではない $\boldsymbol{K}$ 上の線形空間とする $(j = 0, 1, 2, \cdots)$．このとき $\oplus_{j=0}^{\infty} X_j$ は無限次元であることを示せ．

$X$ の双対空間の基底として次のものは非常に有用である．

**定理 21.2**　$X$ を $N$ 次元線形空間で，基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ をもつものとする．このとき，$k = 1, \cdots, N$ に対して

$$f^k\left(\sum_{j=1}^{N} \alpha^j \mathbf{u}_j\right) = \alpha^k$$

と定義すると，$\left\{f^1, \cdots, f^N\right\}$ は $X^*$ の基底になっている．したがって，

$$\dim X^* = \dim X$$

である．

定理の $\left\{f^1, \cdots, f^N\right\}$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ の**双対基底**という．

**証明．** $f_j \in X^*$ は明らか．任意に $f \in X^*$ をとる．このとき，$\mathbf{x} = \sum_{k=1}^{N} \alpha^k \mathbf{u}_k \in X$ に対して，

$$f(\mathbf{x}) = f\left(\sum_{k=1}^{N} \alpha^k \mathbf{u}_k\right) = \sum_{k=1}^{N} f(\mathbf{u}_k) f^k(\mathbf{x}) = \left(\sum_{k=1}^{N} f(\mathbf{u}_k) f^k\right)(\mathbf{x})$$

より $f = \sum_{k=1}^{N} f(\mathbf{u}_k) f^k$ が成り立つ．またもし $\sum_{j=1}^{N} \alpha_j f^j = 0$ であれば，

$$\alpha_k = \sum_{j=1}^{N} \alpha_j f^j(\mathbf{u}_k) = 0.$$

■

二つの線形空間が同形であるという概念を導入しておく．線形空間の理論ではしばしば同形な線形空間は同一視するという立場がとられることがある．

**定義 21.8** $X, Y$ を $\boldsymbol{K}$ 上の線形空間とする．$X$ から $Y$ への線形写像で全単射であるものを**同形写像**という．$X$ から $Y$ への同形写像が存在するとき，$X$ と $Y$ は**同形**であるといい，

$$X \simeq Y$$

と表す．

後の章でテンソル積の意味での同形や多元環の意味での同形などの概念が現れるが，それらと区別を要する際には，本書では同形，同形写像をそれぞれ**線形空間の意味での同形**，**線形空間の意味での同形写像**という．

同形に関する次の性質は基本的なものである．証明は容易なので，読者に委ねる．

**定理 21.3** (1) $X, Y$ を $\boldsymbol{K}$ 上の線形空間とし，$T$ を $X$ から $Y$ への同形写像とする．このとき $T^{-1}$ は $Y$ から $X$ への同形写像である．

(2) $X, Y$ を $\boldsymbol{K}$ 上の線形空間とすると，

$$X \simeq X,$$
$$X \simeq Y \Longrightarrow Y \simeq X,$$
$$X \simeq Y, Y \simeq Z \Longrightarrow X \simeq Z.$$

**定理 21.4** $X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．$\dim X = \dim Y$ とする．このとき $X \simeq Y$ である．逆に $X \simeq Y$ ならば $\dim X = \dim Y$ である．

**証明.** $N = \dim X = \dim Y$ とし，$\{\mathbf{u}_j\}_{j=1}^N$, $\{\mathbf{v}_j\}_{j=1}^N$ をそれぞれ $X, Y$ の基底とする．$\mathbf{x} = \sum_{j=1}^N \alpha_j \mathbf{u}_j \in X$ に対して

$$T(\mathbf{x}) = \sum_{j=1}^N \alpha_j \mathbf{v}_j$$

と定義すると，これが同形写像になっている．逆に $X$ から $Y$ へのある同形写像 $T$ が存在したとする．$\{\mathbf{u}_j\}_{j=1}^N$ を $X$ の基底とし，$\mathbf{v}_j = T(\mathbf{u}_j)$ とおく．$\{\mathbf{v}_j\}_{j=1}^N$ が $Y$ の基底であることを示す．任意の $\mathbf{y} \in Y$ に対して，$T^{-1}(\mathbf{y}) = \mathbf{x}$ とおく．$\mathbf{x} = \sum_{j=1}^N \alpha_j \mathbf{u}_j$ と表せるから，

$$\mathbf{y} = T(\mathbf{x}) = \sum_{j=1}^{N} \alpha_j T(\mathbf{u}_j) = \sum_{j=1}^{N} \alpha_j \mathbf{v}_j$$

である．また $\sum_{j=1}^{N} \alpha_j \mathbf{v}_j = \mathbf{0}$ ならば

$$\sum_{j=1}^{N} \alpha_j \mathbf{u}_j = \sum_{j=1}^{N} \alpha_j T^{-1}(\mathbf{v}_j) = T^{-1}\left(\sum_{j=1}^{N} \alpha_j \mathbf{v}_j\right) = T^{-1}(\mathbf{0}) = \mathbf{0}$$

が成り立つので，$\mathbf{u}_1, \cdots, \mathbf{u}_N$ が線形独立であることより $\alpha_1 = \cdots = \alpha_N = 0$ となる．■

後にしばしば利用する重要な同形写像の例を挙げる．

**例 21.7**　$T \in L(\boldsymbol{K}^N; \boldsymbol{K}^M)$ に対して，ある $A_T \in \boldsymbol{K}^{M \times N}$ で

$$T(\mathbf{x}) = A_T \mathbf{x}, \quad \mathbf{x} \in \boldsymbol{K}^M$$

をみたすものがただ一つ存在する．$\Phi(T) = A_T$ とすると，$\Phi$ は $L(\boldsymbol{K}^N; \boldsymbol{K}^M)$ から $\boldsymbol{K}^{M \times N}$ への同形写像になっている．

**例 21.8**　例 21.6 において，$\mathbf{x} \in X_j$ に対して，

$$\iota(\mathbf{x})_k = \begin{cases} \mathbf{x}, & k = j \\ \mathbf{0}, & k \neq j \end{cases}$$

とし，$\iota(\mathbf{x}) = \{\iota(\mathbf{x})_k\}_{k=0}^{\infty}$ とすると，$\mathbf{x}$ に $\iota(\mathbf{x})$ を対応させる写像 $\iota$ は $X_j$ から $X_j^0$ への同形写像になっている．$\iota$ を $X_j$ の $\oplus_{k=0}^{\infty} X_k$ への**自然な埋め込み**という．この同形写像 $\iota$ により無限直和 $\oplus_{k=0}^{\infty} X_k$ において $X_k$ と $X_k^0 \subset \oplus_{k=0}^{\infty} X_k$ とを同一視して，$X_k$ を $\oplus_{k=0}^{\infty} X_k$ の部分空間とみなすこともある．

次の定理も容易に示せる．

**定理 21.5**　$X$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．$\mathbf{a} \in X$ に対して，

$$\tau_{\mathbf{a}}(f) = f(\mathbf{a}), \quad f \in X^*$$

と定義すると，$\tau_{\mathbf{a}} \in L(X^*; \boldsymbol{K}) = (X^*)^*$ である．このとき，$\mathbf{a} \in X$ に $\tau_{\mathbf{a}}$ を対応させる写像 $\tau$ は $X$ から $(X^*)^*$ への同形写像になっている．$\tau$ を $X$ から $(X^*)^*$ への**自然な同形写像**という．

$(X^*)^*$ は $X$ と同じ次元をもっているから，この間に同形写像が存在する．その

中でも $\tau$ を特に自然な同形写像というのは次のことによる．

**命題 21.6**　$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とし，$\left\{f^1, \cdots, f^N\right\}$ をその双対基底とする．$\{\tau_1, \cdots, \tau_N\}$ を $\left\{f^1, \cdots, f^N\right\}$ の双対基底とすると，

$$\tau_j = \tau_{\mathbf{u}_j}, \quad j = 1, \cdots, N$$

である．

**証明.** $\tau_{\mathbf{u}_j}\left(f^k\right) = f^k(\mathbf{u}_j) = \delta_j^k = \tau_j\left(f^k\right)$ より明らかである． ■

以下ではこの同形写像 $\tau$ により $(X^*)^*$ と $X$ は同一視して考える．

## 21.2　線形写像と行列

本節をとおして $X$ を $\boldsymbol{K}$ 上の有限次元の線形空間とし，$N = \dim X$ とする．$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とし，$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ とおく．このとき，$\mathbf{x} \in X$ に対して，ある $N$ 個のスカラー $x^1, \ldots, x^N \in \boldsymbol{K}$ で

$$\mathbf{x} = \sum_{j=1}^{N} x^j \mathbf{u}_j$$

となるものがただ一組存在する．ここで添え字を上につけているが，これは後に解説するテンソル解析で便利な記法となる．今後 $\mathbf{u}_1, \cdots, \mathbf{u}_N$ に対する線形結合の係数の添え字は上側に付けることにする．

$$\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix}$$

と表し，これを **$\mathbf{x}$ の基底 $\mathcal{U}$ による表示**，あるいは **$\mathbf{x}$ の $\mathcal{U}$-座標**という．$\mathbf{x}_{\mathcal{U}}$ を $(\mathbf{x})_{\mathcal{U}}$ とも表す．

$$X_{\mathcal{U}}(\mathbf{x}) = \mathbf{x}_{\mathcal{U}}, \quad \mathbf{x} \in X$$

により $X$ から $\boldsymbol{K}^N$ への写像を定義し，本書ではこれを **$X$ の $\mathcal{U}$-座標写像**ということにする．次のことが成り立つ．

**定理 21.7**　$X_{\mathcal{U}}$ は $X$ から $\boldsymbol{K}^N$ への同形写像である．

**証明.** まず線形性を証明する．$\mathbf{x} = \sum_{j=1}^{N} x^j \mathbf{u}_j$, $\mathbf{y} = \sum_{j=1}^{N} y^j \mathbf{u}_j$ とする．$\alpha, \beta \in \boldsymbol{K}$

とする．このとき，

$$\alpha \mathbf{x} + \beta \mathbf{y} = \sum_{j=1}^{N} \left( \alpha x^j + \beta y^j \right) \mathbf{u}_j$$

より，$X_{\mathcal{U}}$ は線形である．また $\mathbf{u}_1, \cdots, \mathbf{u}_N$ が基底であることより，$X_{\mathcal{U}}$ が全射であることがわかる．さらに $\mathbf{u}_1, \cdots, \mathbf{u}_N$ の線形独立性から，単射であることも得られる．■

この定理の系として次のことが示せる．

**定理 21.8**　$Z$ を $X$ の非自明な部分空間とする．このとき $Z$ には基底が存在する．また $\{\mathbf{z}_1, \cdots, \mathbf{z}_r\}$ が $Z$ の基底で，$r < N$ であれば，$\mathbf{z}_{r+1}, \cdots, \mathbf{z}_N \in X$ で，$\{\mathbf{z}_1, \cdots, \mathbf{z}_N\}$ が $X$ の基底になるようなものが存在する．

**証明.** $X_{\mathcal{U}}$ を $X$ の $\mathcal{U}$-座標写像とし，

$$X_{\mathcal{U}}(Z) = \{X_{\mathcal{U}}(\mathbf{z}) : \mathbf{z} \in Z\}$$

とおく．定理 21.7 より明らかに $X_{\mathcal{U}}(Z)$ は $\boldsymbol{K}^N$ の非自明な部分空間である．したがって $X_{\mathcal{U}}(Z)$ の基底が存在する (定理 9.14)．$\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ を $X_{\mathcal{U}}(Z)$ の一組の基底とする．このとき，$\mathbf{w}_j = X_{\mathcal{U}}^{-1}(\mathbf{a}_j), j = 1, \cdots, r$ とすると，$\{\mathbf{w}_1, \cdots, \mathbf{w}_r\}$ は $Z$ の基底になっている．

いま $\{\mathbf{z}_1, \cdots, \mathbf{z}_r\}$ は $Z$ の任意の基底とする．このとき $\{X_{\mathcal{U}}(\mathbf{z}_1), \cdots, X_{\mathcal{U}}(\mathbf{z}_r)\}$ は $X_{\mathcal{U}}(Z)$ の基底である．したがって，$r < N$ の場合は，定理 9.19(あるいは定理 9.8 参照) より，この基底に $N - r$ 個のベクトル $\mathbf{a}_{r+1}, \cdots, \mathbf{a}_N \in \boldsymbol{K}^N$ を加えて，$\boldsymbol{K}^N$ の基底を作ることができる．$\mathbf{z}_j = X_{\mathcal{U}}^{-1}(\mathbf{a}_j), j = r+1, \cdots, N$ とおけば，$\{\mathbf{z}_1, \cdots, \mathbf{z}_N\}$ は $X$ の基底になっている．■

さて，$X, Y$ を有限次元の線形空間であるとする．$N = \dim X$, $M = \dim Y$ とする．$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底，$\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_M\}$ を $Y$ の基底であるとする．$X_{\mathcal{U}}, Y_{\mathcal{V}}$ をそれぞれ $X$ の $\mathcal{U}$-座標写像，$Y$ の $\mathcal{V}$-座標写像とする．$T$ を $X$ から $Y$ への線形写像とするとき，

$$\Phi(T) = Y_{\mathcal{V}} \circ T \circ X_{\mathcal{U}}^{-1}$$

は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像になる．逆に $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への任意の線形写像 $S$ に対しては，

$$\Psi(S) = Y_{\mathcal{V}}^{-1} \circ S \circ X_{\mathcal{U}}$$

により $X$ から $Y$ への線形写像が定義される．明らかに

$$\Psi(\Phi(T)) = T, \quad \Phi(\Psi(S)) = S$$

である．$\Phi$ は $L(X;Y)$ から $L(\boldsymbol{K}^N;\boldsymbol{K}^M)$ への同形写像になっている．そこで $\Phi$ を介して $X$ から $Y$ への線形写像は $\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像とみなすことができる．$\boldsymbol{K}^N$ から $\boldsymbol{K}^M$ への線形写像は $(M,N)$ 行列により表すことができるので，抽象的な設定での線形写像は，適当な基底をとることにより，行列の解析に帰着することができる．

**問題 21.2**　$L(X;Y)$ と $\boldsymbol{K}^{M\times N}$ が同形であることを示せ．

## 21.3　座標変換について

$X$ を $\boldsymbol{K}$ 上の有限次元の線形空間とし，$N = \dim X$ とする．$\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$，$\widetilde{\mathcal{U}} = \{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ を $X$ の基底とする．このとき，

$$\widetilde{\mathbf{u}}_j = \sum_{\mu=1} u_j^{\mu}\, \mathbf{u}_\mu, \quad j = 1, \cdots, N \tag{21.1}$$

をみたすようなスカラー $u_j^\mu$, $j, \mu = 1, \cdots, N$ が一意的に存在する．明らかに行列

$$\left(u_j^\mu\right) = \begin{pmatrix} u_1^1 & \cdots & u_N^1 \\ \vdots & \ddots & \vdots \\ u_1^N & \cdots & u_N^N \end{pmatrix}$$

は正則である．$\left(u_j^\mu\right)$ の逆行列を $\left(\widetilde{u}_j^\mu\right)$ とおく．このとき

$$\sum_{l=1}^{N} u_l^k\, \widetilde{u}_j^l = \sum_{l=1}^{N} \widetilde{u}_l^k\, u_j^l = \delta_j^k \tag{21.2}$$

である．ただし $\delta_j^k$ はクロネッカーのデルタ記号である．これより

$$\mathbf{u}_j = \sum_{k=1} \widetilde{u}_j^k\, \widetilde{\mathbf{u}}_k, \quad j = 1, \cdots, N \tag{21.3}$$

が成り立っていることがわかる．なぜならば

$$\mathbf{u}_j = \sum_{\mu=1}^{N} \left( \sum_{k=1}^{N} \widetilde{u}_j^k\, u_k^\mu \right) \mathbf{u}_\mu = \sum_{k=1}^{N} \widetilde{u}_j^k \left( \sum_{\mu=1}^{N} u_k^\mu\, \mathbf{u}_\mu \right) = \sum_{k=1}^{N} \widetilde{u}_j^k\, \widetilde{\mathbf{u}}_k. \tag{21.4}$$

この関係式をもとにして，$\mathbf{x}_{\mathcal{U}}$ と $\mathbf{x}_{\widetilde{\mathcal{U}}}$ の間の関係式を求めてみよう．

**定理 21.9**　以上の設定において

$$\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix}, \quad \mathbf{x}_{\widetilde{\mathcal{U}}} = \begin{pmatrix} \widetilde{x}^1 \\ \vdots \\ \widetilde{x}^N \end{pmatrix}$$

とすると，この間に次の変換式が成り立つ．

$$\begin{pmatrix} \widetilde{x}^1 \\ \vdots \\ \widetilde{x}^N \end{pmatrix} = \begin{pmatrix} \widetilde{u}_1^1 & \cdots & \widetilde{u}_N^1 \\ \vdots & \ddots & \vdots \\ \widetilde{u}_1^N & \cdots & \widetilde{u}_N^N \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix}, \tag{21.5}$$

$$\begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix} = \begin{pmatrix} u_1^1 & \cdots & u_N^1 \\ \vdots & \ddots & \vdots \\ u_1^N & \cdots & u_N^N \end{pmatrix} \begin{pmatrix} \widetilde{x}^1 \\ \vdots \\ \widetilde{x}^N \end{pmatrix}. \tag{21.6}$$

**証明.** 定義から

$$\sum_{k=1}^{N} \widetilde{x}^k \widetilde{\mathbf{u}}_k = \sum_{j=1}^{j} x^j \mathbf{u}_j$$

である．したがって (21.3) より

$$\sum_{k=1}^{N} \widetilde{x}^k \widetilde{\mathbf{u}}_k = \sum_{j=1}^{j} x^j \mathbf{u}_j = \sum_{j=1}^{N} x^j \sum_{k=1}^{N} \widetilde{u}_j^k \widetilde{\mathbf{u}}_k$$
$$= \sum_{k=1}^{N} \left( \sum_{j=1}^{N} \widetilde{u}_j^k x^j \right) \widetilde{\mathbf{u}}_k.$$

ゆえに $k = 1, \cdots, N$ に対して

$$\widetilde{x}^k = \sum_{j=1}^{N} \widetilde{u}_j^k x^j$$

が成り立っている．これより (21.5) が示された．(21.6) は $\left(u_j^k\right)$ が $\left(\widetilde{u}_j^k\right)$ の逆行列であることによる．■

次に $X^*$ について考えてみよう．$\mathcal{F} = \left\{\mathbf{f}^1, \cdots, \mathbf{f}^N\right\}$ を $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ の双対基底とする ($\mathcal{F}$ は離散フーリエ変換と同じ記号であるが，文脈から混乱は生じないであろう)．また $\widetilde{\mathcal{F}} = \left\{\widetilde{\mathbf{f}}^1, \cdots, \widetilde{\mathbf{f}}^N\right\}$ を $\widetilde{\mathcal{X}} = \{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ の双対基底とする．このとき $\widetilde{\mathcal{F}}$ と $\mathcal{F}$ の間には次の関係式が成り立っている．

$$\mathbf{f}^j = \sum_{k=1}^{N} u_k^j \widetilde{\mathbf{f}}^k, \quad \widetilde{\mathbf{f}}^j = \sum_{k=1}^{N} \widetilde{u}_k^j \mathbf{f}^k \tag{21.7}$$

**証明.** $\widetilde{\mathbf{u}}_\mu = \sum_{\nu=1}^{N} u_k^\nu \mathbf{u}_\nu$ より

$$\mathbf{f}^j(\widetilde{\mathbf{u}}_\mu) = \mathbf{f}^j\left(\sum_{\nu=1}^{N} u_\mu^\nu \mathbf{u}_\nu\right) = \sum_{\nu=1}^{N} u_\mu^\nu \mathbf{f}^j(\mathbf{u}_\nu) = u_\mu^j = \sum_{k=1}^{N} u_k^j \widetilde{\mathbf{f}}^k(\widetilde{\mathbf{u}}_\mu)$$

である．後半は次のようにして得られる．

$$\widetilde{\mathbf{f}}^j = \sum_{\mu=1}^{N} \delta_\mu^j \widetilde{\mathbf{f}}^\mu = \sum_{\mu=1}^{N}\left(\sum_{\nu=1}^{N} \widetilde{u}_\nu^j u_\mu^\nu\right)\widetilde{\mathbf{f}}^\mu = \sum_{\nu=1}^{N} \widetilde{u}_\nu^j \left(\sum_{\mu=1}^{N} u_\mu^\nu \widetilde{\mathbf{f}}^\mu\right)$$
$$= \sum_{\nu=1}^{N} \widetilde{u}_\nu^j \mathbf{f}^\nu.$$

■

さて，任意の $\mathbf{f} \in X^*$ に対して

$$\mathbf{f} = \sum_{k=1}^{N} \alpha_k \mathbf{f}^k = \sum_{k=1}^{N} \widetilde{\alpha}_k \widetilde{\mathbf{f}}^k$$

をみたす $\alpha_k, \widetilde{\alpha}_k \in \boldsymbol{K}$, $k = 1, \cdots, N$ が一意的に存在する．このとき，

$$\mathbf{f}_{\mathcal{F}} = \begin{pmatrix} \alpha_1 & \cdots & \alpha_N \end{pmatrix}, \quad \mathbf{f}_{\widetilde{\mathcal{F}}} = \begin{pmatrix} \widetilde{\alpha}_1 & \cdots & \widetilde{\alpha}_N \end{pmatrix}$$

と表し，それぞれ $\mathbf{f}$ の $\mathcal{F}$-座標，$\widetilde{\mathcal{F}}$-座標という．$S_{\mathcal{F}}(\mathbf{f}) = \mathbf{f}_{\mathcal{F}}$ とすると，$S_{\mathcal{F}}$ は $X^*$ から $\boldsymbol{K}^{1\times N}$ への同形写像になっている．$\mathbf{f}_{\mathcal{F}}$ と $\mathbf{f}_{\widetilde{\mathcal{F}}}$ の間の関係式を求める．

**定理 21.10** 以上の設定において次のことが成り立つ．

$$\begin{pmatrix} \widetilde{\alpha}_1 & \cdots & \widetilde{\alpha}_N \end{pmatrix} = \begin{pmatrix} \alpha_1 & \cdots & \alpha_N \end{pmatrix} \begin{pmatrix} u_1^1 & \cdots & u_N^1 \\ \vdots & \ddots & \vdots \\ u_1^N & \cdots & u_N^N \end{pmatrix}, \tag{21.8}$$

$$\begin{pmatrix} \alpha_1 & \cdots & \alpha_N \end{pmatrix} = \begin{pmatrix} \widetilde{\alpha}_1 & \cdots & \widetilde{\alpha}_N \end{pmatrix} \begin{pmatrix} \widetilde{u}_1^1 & \cdots & \widetilde{u}_N^1 \\ \vdots & \ddots & \vdots \\ \widetilde{u}_1^N & \cdots & \widetilde{u}_N^N \end{pmatrix}. \tag{21.9}$$

**証明.** $\mathbf{f} = \sum_{k=1}^{N} \alpha_k \mathbf{f}^k = \sum_{k=1}^{N} \widetilde{\alpha}_k \widetilde{\mathbf{f}}^k$ であるから，(21.7) より

$$\sum_{k=1}^{N}\widetilde{\alpha}_k\widetilde{\mathbf{f}}^k=\sum_{j=1}^{N}\alpha_j f^j=\sum_{j=1}^{N}\alpha_j\sum_{k=1}^{N}u_k^j\widetilde{\mathbf{f}}^k=\sum_{k=1}^{N}\left(\sum_{j=1}^{N}u_k^j\alpha_j\right)\widetilde{\mathbf{f}}^k$$

である．したがって $\widetilde{\alpha}_k=\sum_{j=1}^{N}u_k^j\alpha_j$ となり，(21.8) が証明された．後半の主張は $(u_j^k)$ が $(\widetilde{u}_j^k)$ の逆行列であることによる．■

**注意 21.2**　(21.8), (21.9) の転置をとれば次式が得られる．

$$\begin{pmatrix}\widetilde{\alpha}_1\\ \vdots\\ \widetilde{\alpha}_1\end{pmatrix}=\begin{pmatrix}u_1^1 & \cdots & u_1^N\\ \vdots & \ddots & \vdots\\ u_N^1 & \cdots & u_N^N\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_1\end{pmatrix},$$

$$\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_1\end{pmatrix}=\begin{pmatrix}\widetilde{u}_1^1 & \cdots & \widetilde{u}_1^N\\ \vdots & \ddots & \vdots\\ \widetilde{u}_N^1 & \cdots & \widetilde{u}_N^N\end{pmatrix}\begin{pmatrix}\widetilde{\alpha}_1\\ \vdots\\ \widetilde{\alpha}_1\end{pmatrix}.$$

定理 21.9 からわかることは，$\mathcal{U}$-座標 $\mathbf{x}_{\mathcal{U}}$ を $\widetilde{\mathcal{U}}$-座標 $\mathbf{x}_{\widetilde{\mathcal{U}}}$ に変換する行列が，$\mathcal{U}=\{\mathbf{u}_1,\cdots,\mathbf{u}_N\}$ を $\widetilde{\mathcal{U}}=\{\widetilde{\mathbf{u}}_1,\cdots,\widetilde{\mathbf{u}}_N\}$ に変換する行列の逆行列になっていることである．一方，定理 21.10 より $\{\mathbf{f}^1,\cdots,\mathbf{f}^N\}$ に関する座標を $\{\widetilde{\mathbf{f}}^1,\cdots,\widetilde{\mathbf{f}}^N\}$ に関する座標に変換する行列が，$\mathcal{U}$ を $\widetilde{\mathcal{U}}$ に変換する行列そのものになっていることがわかることから，$X$ の要素を**反変ベクトル**といい，$X^*$ の要素を**共変ベクトル**という．

## 21.4　内積

内積を一般の線形空間に定義する．

**定義 21.9**　$X$ を $\boldsymbol{K}$ 上の線形空間とする．$X$ の任意の元 $\mathbf{x},\mathbf{y}$ に対して $\boldsymbol{K}$ の元 $\langle\mathbf{x},\mathbf{y}\rangle$ が対応していて，次の (1), (2), (3) をみたすとき，$\langle\cdot,\cdot\rangle$ を $X$ 上の**内積**という．

(1) (正値性)　$\langle\mathbf{x},\mathbf{x}\rangle\geq 0\ (\mathbf{x}\in X)$; $\langle\mathbf{x},\mathbf{x}\rangle=0\Longleftrightarrow\mathbf{x}=\mathbf{0}$.

(2) (共役対称性)　$\langle\mathbf{x},\mathbf{y}\rangle=\overline{\langle\mathbf{y},\mathbf{x}\rangle}\quad(\mathbf{x},\mathbf{y}\in X)$.

(3) (線形性)　$\langle\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{z}\rangle=\alpha\langle\mathbf{x},\mathbf{z}\rangle+\beta\langle\mathbf{y},\mathbf{z}\rangle\quad(\mathbf{x},\mathbf{y},\mathbf{z}\in X,\ \alpha,\beta\in\boldsymbol{K})$.

$\boldsymbol{K}$ 上の線形空間 $X$ 上に内積 $\langle\cdot,\cdot\rangle$ が定義されているとき，$X$ と内積との組 $(X, \langle\cdot,\cdot\rangle)$ を $\boldsymbol{K}$ 上の**内積空間**という．

$X = \boldsymbol{K}^N$ の場合に定義 10.1 で定めた内積，すなわち $\mathbf{x} = (x_j)$ , $\mathbf{y} = (y_j) \in \boldsymbol{K}$ に対して

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{N} x_j \, \overline{y_j}$$

と定義したものは定義 21.9 の意味で $\boldsymbol{K}^N$ 上の内積になっている．この内積を $\boldsymbol{K}^N$ の**ユークリッド内積**という．今後は混乱のないようにユークリッド内積を $\langle\cdot,\cdot\rangle_{\boldsymbol{K}^N}$ と表す．また特に言明しない限り，$\boldsymbol{K}^N$ はユークリッド内積 $\langle\cdot,\cdot\rangle_{\boldsymbol{K}^N}$ が定義された内積空間であるとする．

一つの線形空間上に複数の内積が定義されうる．例として次のようなものがある．

**例 21.9** $A$ を $N$ 次のエルミート行列で，その固有値はすべて正であるとする．すなわち正定値行列であるとする．$\mathbf{x}, \mathbf{y} \in \boldsymbol{K}^N$ に対して，

$$\langle \mathbf{x}, \mathbf{y} \rangle_A = \langle A\mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^N}$$

とする．このとき $\langle\cdot,\cdot\rangle_A$ は $\boldsymbol{K}^N$ 上の内積である．

**証明．** あるユニタリー行列 $U$ により，$U^*AU$ =diag$(\lambda_1, \cdots, \lambda_N)$ と対角化する．仮定より $\lambda_j > 0, j = 1, \cdots, N$ である．$U^*\mathbf{x} = (z_j)$ とおく．このとき，定理 13.16 より

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{x} \rangle_A &= \langle A\mathbf{x}, \mathbf{x} \rangle_{\boldsymbol{K}^N} = \langle U^*AUU^*\mathbf{x}, U^*\mathbf{x} \rangle_{\boldsymbol{K}^N} \\
&= \langle \mathrm{diag}\,(\lambda_1, \cdots, \lambda_N)\, U^*\mathbf{x}, U^*\mathbf{x} \rangle_{\boldsymbol{K}^N} \\
&= \sum_{j=1}^{N} \lambda_j \, |z_j|^2
\end{aligned}$$

である．したがって内積の条件 (1) が成り立つ．また

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle_A &= \langle A\mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{K}^N} = \langle \mathbf{x}, A\mathbf{y} \rangle_{\boldsymbol{K}^N} = \overline{\langle A\mathbf{y}, \mathbf{x} \rangle_{\boldsymbol{K}^N}} \\
&= \overline{\langle \mathbf{y}, \mathbf{x} \rangle_A}
\end{aligned}$$

より条件 (2) が成立する．(3) は明らかである．■

**定理 21.11** $(X, \langle\cdot,\cdot\rangle)$ を $\boldsymbol{K}$ 上の内積空間とする．$\mathbf{x} \in X$ に対して $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ とする．このとき，$\|\cdot\|$ は $\mathbf{x}, \mathbf{y} \in X$, $\alpha \in \boldsymbol{K}$ に対して

$$\|\mathbf{x}\| \geq 0\,;\ \|\mathbf{x}\| = 0 \Longleftrightarrow \mathbf{x} = 0, \tag{21.10}$$
$$\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|, \tag{21.11}$$
$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \tag{21.12}$$
$$|\langle \mathbf{x},\mathbf{y}\rangle| \leq \|\mathbf{x}\|\,\|\mathbf{y}\| \tag{21.13}$$

をみたす．(21.12) を**三角不等式**，(21.13) を**シュヴァルツの不等式**という．

**証明．** (21.10), (21.11) は内積の定義から明らかである．シュヴァルツの不等式を証明する．$\langle \mathbf{x},\mathbf{y}\rangle \neq 0$ の場合を示す．任意の実数 $t$ に対して，

$$\begin{aligned} 0 &\leq \langle \mathbf{x} + t\,\langle \mathbf{x},\mathbf{y}\rangle\,\mathbf{y}, \mathbf{x} + t\,\langle \mathbf{x},\mathbf{y}\rangle\,\mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + 2t\,|\langle \mathbf{x},\mathbf{y}\rangle|^2 + t^2\,|\langle \mathbf{x},\mathbf{y}\rangle|^2\,\|\mathbf{y}\|^2 \end{aligned}$$

である．これを $t$ の 2 次式とみて，判別式をとると “判別式 $\leq 0$” であるから

$$4\,|\langle \mathbf{x},\mathbf{y}\rangle|^4 - 4\,|\langle \mathbf{x},\mathbf{y}\rangle|^2\,\|\mathbf{x}\|^2\,\|\mathbf{y}\|^2 \leq 0$$

である．ゆえにシュヴァルツの不等式が成り立つ．定理 10.4 と同様にして (21.12) が証明される．■

**定義 21. 10**　定理 21.11 で定義された $\|\cdot\|$ を**内積** $\langle\cdot,\cdot\rangle$ **から誘導されるノルム**という．

$$d\,(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$$

を**内積** $\langle\cdot,\cdot\rangle$ **から誘導される距離**という．特に $\boldsymbol{K}^N$ 上のユークリッド内積から誘導されるノルムを**ユークリッド・ノルム**，またユークリッド内積から誘導される距離を**ユークリッド距離**という．

**定義 21. 11**　一般に線形空間 $X$ 上の関数 $\|\cdot\|$ で条件 (21.10), (21.11), (21.12) をみたすものを $X$ 上の**ノルム**という．

**例 21. 10**　$X \in \boldsymbol{R}^{N\times p}$ をデータ行列とし，$S_X$ をその分散共分散行列とする．すなわち $S_X = N^{-1}\widetilde{X}^T\,\widetilde{X}$ とする．いま $S_X$ が正定値であるとする．このとき $S_X^{-1}$ も正定値であるから $\boldsymbol{R}^p$ 上に内積 $\langle\cdot,\cdot\rangle_{S_X^{-1}}$ が定義され，それから誘導される距離を考えることができる．その距離を $d_{S_X}(\cdot,\cdot)$ により表す．$d_{S_X}(\cdot,\cdot)$ を**マハラノビス距離**という．$\overline{\mathbf{x}}$ を $X$ の平均ベクトルとする．このとき $c > 0$ に対して，

$$\{\mathbf{x} : \mathbf{x} \in \boldsymbol{R}^p, d_{S_X}(\mathbf{x},\overline{\mathbf{x}}) = c\}$$

を大きさ $c$ の**集中楕円**という．マハラノビス距離は多変量解析の判別分析などに用いられる．

**問題 21.3** $X \in \boldsymbol{R}^{N\times p}$ をデータ行列とし，その分散共分散行列は正定値であるとする．$d(\cdot,\cdot)$ を $\boldsymbol{R}^p$ 上のユークリッド距離とする．このときある正則行列 $T \in \boldsymbol{R}^{p\times p}$ が存在し，任意の $\mathbf{x}, \mathbf{x}' \in \boldsymbol{R}^p$ に対して，

$$d(T\mathbf{x}, T\mathbf{x}') = d_{S_X}(\mathbf{x}, \mathbf{x}')$$

が成り立つ．

$X$ を有限次元の線形空間とし，これに内積 $\langle\cdot,\cdot\rangle$ が定義されているとする．有限次元の定義から $X$ の基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ が存在するが，特に

$$\langle \mathbf{u}_j, \mathbf{u}_k\rangle = \delta_{jk}$$

をみたすとき，$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の $\langle\cdot,\cdot\rangle$ に関する**正規直交基底**という．この場合にもシュミットの直交化法はそのまま適用することができ，$X$ の基底から正規直交基底をつくることができる．したがって，有限次元の内積空間には必ず正規直交基底が存在する．

**定理 21.12** 任意の $\mathbf{a} \in X$ に対して，

$$\iota_{\mathbf{a}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{a}\rangle, \quad \mathbf{x} \in X$$

とすると，$\iota_{\mathbf{a}} \in X^*$ である．逆に，任意の $\mathbf{f} \in X^*$ に対して，ある $\mathbf{a} \in X$ で $\mathbf{f} = \iota_{\mathbf{a}}$ となるものが一意的に存在する．

**証明.** 前半の主張は明らかなので，後半を証明する．

$N = \dim X$ とし，$\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ をその正規直交基底とする．$\mathbf{f}(\mathbf{x}_j) = \alpha_j$ とおく．$\mathbf{a} = \sum_{j=1}^{N} \overline{\alpha_j}\mathbf{x}_j$ と定義する．任意の $\mathbf{x} \in X$ をとり $\mathbf{x} = \sum_{j=1}^{N} \beta_j \mathbf{x}_j$ とする．このとき

$$\mathbf{f}(\mathbf{x}) = \sum_{j=1}^{N} \beta_j \alpha_j = \langle \mathbf{x}, \mathbf{a}\rangle = \iota_{\mathbf{a}}(\mathbf{x})$$

が成り立っている．次に $\mathbf{a}$ の一意性を示す．もしある $\mathbf{a}$, $\mathbf{a}'$ が存在し，任意の $\mathbf{x} \in X$ に対して，

$$\mathbf{f}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{a}\rangle = \langle \mathbf{x}, \mathbf{a}'\rangle$$

となるならば，$\langle \mathbf{x}, \mathbf{a} - \mathbf{a}'\rangle = 0$ である．特に $\mathbf{x} = \mathbf{a} - \mathbf{a}'$ を考えれば，$\|\mathbf{a} - \mathbf{a}'\| = 0$

が得られる．■

**系 21. 13**　$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の正規直交基底とし，$\{\mathbf{f}^1, \cdots, \mathbf{f}^N\}$ をその双対基底とする．このとき

$$\iota_{\mathbf{u}_j} = \mathbf{f}^j, \quad j = 1, \cdots, N$$

である．

**証明.** 次のことから明らかである．

$$\iota_{\mathbf{u}_j}(\mathbf{u}_k) = \langle \mathbf{u}_k, \mathbf{u}_j \rangle = \delta_{jk}.$$

■

$\mathbf{f} \in X^*$ に対して，

$$\|\mathbf{f}\|_{X^*} = \sup\{|f(\mathbf{x})| : \mathbf{x} \in X,\ \|\mathbf{x}\| \le 1\}$$

とする．このとき，次のことが成り立つ．

**定理 21. 14**　$\mathbf{f}, \mathbf{g} \in X^*$, $\alpha \in \boldsymbol{K}$ に対して，

(1)　$\|\mathbf{f}\|_{X^*} \ge 0.$

(2)　$\|\mathbf{f}\|_{X^*} = 0 \Longleftrightarrow \mathbf{f} = 0.$

(3)　$\|\alpha\mathbf{f}\|_{X^*} = |\alpha|\,\|\mathbf{f}\|_{X^*}.$

(4)　$\|\mathbf{f}+\mathbf{g}\|_{X^*} \le \|\mathbf{f}\|_{X^*} + \|\mathbf{g}\|_{X^*}.$

証明は容易なので省略する．$\|\cdot\|_{X^*}$ を $X^*$ 上の**汎関数ノルム**という．

**系 21. 15**　$\iota$ により $\mathbf{a} \in X$ に $\iota_{\mathbf{a}} \in X^*$ を対応される写像を表す．このとき，$\iota$ は $X$ から $X^*$ への同形写像になる．さらに

$$\mathbf{f}(\mathbf{x}) = \left\langle \mathbf{x}, \iota^{-1}(\mathbf{f}) \right\rangle, \quad \mathbf{f} \in X^*, \mathbf{x} \in X,$$
$$\|\mathbf{a}\| = \|\iota_{\mathbf{a}}\|_{X^*}$$

をみたす．

**証明.** 定理 21.12 より，同形写像であることは明らか．シュヴァルツの不等式より

$$|\iota_{\mathbf{a}}(\mathbf{x})| = |\langle \mathbf{x}, \mathbf{a} \rangle| \le \|\mathbf{x}\|\,\|\mathbf{a}\|$$

であるから，$\|\iota_{\mathbf{a}}\|_{X^*} \le \|\mathbf{a}\|$ が成り立つことがわかる．$\mathbf{a} \ne \mathbf{0}$ の場合，$\mathbf{x} = \|\mathbf{a}\|^{-1}\mathbf{a}$

とすると $\|\mathbf{x}\| = 1$ であり，したがって

$$\|\iota_{\mathbf{a}}\|_{X^*} \geq |\langle \mathbf{a}, \mathbf{x} \rangle| = \|\mathbf{a}\|$$

である．残りの主張は明らかである． ■

# 第22章　テンソル積と外積

本章では線形空間，線形写像のテンソル積および外積について述べる．これらは応用上も重要な概念である．

**ポイント22**　線形空間のテンソル積，外積は数学の多くの分野で用いられている．たとえば微分幾何で学ぶ微分形式などが代表例であろう．そのほか，離散フーリエ解析や離散ウェーブレット解析を画像のような2次元配列からなるデータに適用しようとする際，テンソル積の考え方を用いることができる．

## 22.1　線形空間のテンソル積

この節では線形空間のテンソル積の定義をする．前章で定めたように，$\boldsymbol{K}$ 上の線形空間 $X_1, \cdots, X_n$ に対して，$X_1 \times \cdots \times X_n$ から $W$ への $n$ 重線形写像全体のなす集合を $L(X_1, \cdots, X_n; W)$ と表す．

**定義 22.1**　$X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の線形空間とする．$\boldsymbol{K}$ 上の線形空間 $V$ と $\kappa \in L(X_1, \cdots, X_n; V)$ が次の (1), (2) をみたすとき，組 $(V, \kappa)$ を $X_1, \cdots, X_n$ の**テンソル積**という．

(1)　$\kappa\,(X_1, \cdots, X_n) = \{\kappa\,(\mathbf{x}_1, \cdots, \mathbf{x}_n) : \mathbf{x}_j \in X_j,\ j = 1, \cdots, n\}$ とすると，

$$\operatorname{span} \kappa\,(X_1, \cdots, X_n) = V.$$

(2)　$\boldsymbol{K}$ 上の任意の線形空間 $W$ と任意の $\Phi \in L(X_1, \cdots, X_n; W)$ に対して，ある $T \in L(V; W)$ で

$$\Phi\,(\mathbf{x}_1, \cdots, \mathbf{x}_n) = T\,(\kappa\,(\mathbf{x}_1, \cdots, \mathbf{x}_n)), \quad \mathbf{x}_j \in X_j,\ j = 1, \cdots, n \tag{22.1}$$

となるものが存在する．

$$\begin{array}{ccc} X_1 \times \cdots \times X_n & \underset{\kappa}{\longrightarrow} & V \\ & \underset{\Phi}{\searrow} & \downarrow T \\ & & W \end{array}$$

有限次元の場合にテンソル積が存在することは後で証明するが，一般にはテンソル積は一意的ではない．そこで二つのテンソル積の間に次のような同形の定義を導入し，同形なテンソル積はしばしば同一視する．また補題 22.2 で示されるように，定義 22.1 (2) の $T$ は一意的に定まる．

**定義 22. 2**　$X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の線形空間とし, $(V, \kappa)$ , $(V', \kappa')$ を $X_1, \cdots, X_n$ のテンソル積とする．このとき，$V$ から $V'$ への同形写像 $\chi$ で，

$$\kappa'(\mathbf{x}_1, \cdots, \mathbf{x}_n) = \chi(\kappa(\mathbf{x}_1, \cdots, \mathbf{x}_n)), \quad \mathbf{x}_j \in X_j,\ j = 1, \cdots, n \tag{22.2}$$

をみたすものが存在するとき，これをテンソル積の意味での**同形写像**といい，$(V, \tau)$ , $(V', \tau')$ はテンソル積の意味で**同形**であるという．またこのことを単に $(V, \kappa) \approx (V', \kappa')$ あるいは $V \approx V'$ と表す．

$$\begin{array}{ccc} X_1 \times \cdots \times X_n & \underset{\kappa}{\longrightarrow} & V \\ & \underset{\kappa'}{\searrow} & \updownarrow \chi \\ & & V' \end{array}$$

前後の文脈から線形空間の意味での同形あるいは同形写像と混乱の恐れがない場合，テンソル積の意味での同形，テンソル積の意味での同形写像を単に同形，同形写像と呼ぶ．

テンソル積の単純な例として，$\boldsymbol{K}^M$ と $\boldsymbol{K}^N$ のテンソル積の例を示しておこう．

**例 22. 1**　$V = \boldsymbol{K}^{M \times N}$ とする．$\boldsymbol{K}^M$ の標準基底 $\left\{\mathbf{e}_j^{(M)}\right\}_{j=1}^{M}$ と $\boldsymbol{K}^N$ の標準基底 $\left\{\mathbf{e}_k^{(N)}\right\}_{k=1}^{N}$ を考える．$\mathbf{x} = \sum_{j=1}^{M} x_j \mathbf{e}_j^{(M)} \in \boldsymbol{K}^M$ と $\mathbf{y} = \sum_{k=1}^{N} y_k \mathbf{e}_k^{(N)}$ に対して，

$$\kappa(\mathbf{x}, \mathbf{y}) = \mathbf{x}\,\mathbf{y}^T = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_N \\ \vdots & \ddots & \vdots \\ x_M y_1 & \cdots & x_M y_N \end{pmatrix} \in V$$

と定義する．このとき，$\left(\boldsymbol{K}^{M \times N}, \kappa\right)$ は $\boldsymbol{K}^M$ と $\boldsymbol{K}^N$ のテンソル積になっている．

特に $\boldsymbol{K}^1$ と $\boldsymbol{K}^1$ のテンソル積として $(\boldsymbol{K}^1, \kappa)$，ただし

$$\kappa(x, y) = xy, \quad x, y \in \boldsymbol{K}^1$$

をテンソル積ととることができる．

**証明.** 明らかに $\kappa \in L\left(\boldsymbol{K}^M, \boldsymbol{K}^N; \boldsymbol{K}^{M \times N}\right)$ である．$\kappa\left(\mathbf{e}_j^{(M)}, \mathbf{e}_k^{(N)}\right)$ は $(j, k)$ 成分のみが 1 で，残りは 0 からなる $M \times N$ 行列である．このような行列の全体は $\boldsymbol{K}^{M \times N}$ の基底をなしている．ゆえに定義 22.1 (1) が示せた．定義 22.1 (2) が成り立つことを示す．$\boldsymbol{K}$ 上の任意の線形空間 $W$ と任意の $\Phi \in L(\boldsymbol{K}^M, \boldsymbol{K}^N; W)$ をとる．$\mathbf{e}_{jk} = \kappa\left(\mathbf{e}_j^{(M)}, \mathbf{e}_k^{(N)}\right)$ とおくと，任意の $\mathbf{a} \in V$ は

$$\mathbf{a} = \sum_{j=1}^{M} \sum_{k=1}^{N} a_{jk} \mathbf{e}_{jk}$$

と一意的に表せている．そこで

$$T(\mathbf{a}) = \sum_{j=1}^{M} \sum_{k=1}^{N} a_{jk} \Phi\left(\mathbf{e}_j^{(M)}, \mathbf{e}_k^{(N)}\right)$$

と定義する．このとき容易に $T \in L(V; W)$ であることが示せる．さらに任意の $\mathbf{x} = \sum_{j=1}^{M} x_j \mathbf{e}_j^{(M)}$ と $\mathbf{y} = \sum_{k=1}^{N} y_k \mathbf{e}_k^{(N)}$ に対して，

$$\kappa(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{M} \sum_{k=1}^{N} x_j y_k \mathbf{e}_{jk}$$

であるから，

$$\begin{aligned} T(\kappa(\mathbf{x}, \mathbf{y})) &= \sum_{j=1}^{M} \sum_{k=1}^{N} x_j y_k \Phi\left(\mathbf{e}_j^{(M)}, \mathbf{e}_k^{(N)}\right) = \Phi\left(\sum_{j=1}^{M} x_j \mathbf{e}_j^{(M)}, \sum_{k=1}^{N} y_k \mathbf{e}_k^{(N)}\right) \\ &= \Phi(\mathbf{x}, \mathbf{y}) \end{aligned}$$

である．よって定義 22.1 (2) が示せた．■

次の定理を証明する．

**定理 22.1**　$X_1, \cdots, X_n$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．このとき $X_1, \cdots, X_n$ のテンソル積が存在する．また $X_1, \cdots, X_n$ のテンソル積はみな互いにテンソル積の意味で同形である．

定理 22.1 の証明のため，次の補題を準備する．

**補題 22.2** 定義 22.1 の (2) の写像 $T$ は $W$ と $\Phi$ が定まればそれに対して一意的に定まる．したがって定義 22.2 の同形写像 $\chi$ は $(V,\kappa)$, $(V',\kappa')$ に対して一意的に定まる．

**証明.** もし $T'$ が (2) をみたす別の写像であるとすると,

$$T\circ\kappa = T'\circ\kappa$$

であるが，(1) の条件と $T, T'$ の線形性より $T=T'$ が得られる． ■

**補題 22.3** $\mathbf{x}_j\in X_j$, $j=1,\cdots,n$ をとり固定する．このとき

$$\tau_{\mathbf{x}_1,\cdots,\mathbf{x}_n}(\mathbf{f}_1,\cdots,\mathbf{f}_n)=\mathbf{f}_1(\mathbf{x}_1)\cdots\mathbf{f}_n(\mathbf{x}_n),\quad \mathbf{f}_j\in X_j^*,\ j=1,\cdots,n$$

と定義すると，$\tau_{\mathbf{x}_1,\cdots,\mathbf{x}_n}\in L(X_1^*,\cdots,X_n^*;\boldsymbol{K})$ である．

証明は容易なので省略する．

**定理 22.1 の証明.** 以下では $n=2$ の場合を証明する．一般の $n$ の場合も，記号が煩雑になるが同様にして証明できる．$X=X_1$, $Y=X_2$ とおく．$V=L(X^*,Y^*;\boldsymbol{K})$ とおく．$\mathbf{x}\in X$, $\mathbf{y}\in Y$ に対して

$$\kappa(\mathbf{x},\mathbf{y})=\tau_{\mathbf{x},\mathbf{y}}$$

と定義すると，補題 22.3 より，$\kappa\in L(X,Y;V)$ である．これが定義 22.1 (1), (2) をみたすことを証明する．$X$ の基底 $\{\mathbf{u}_j\}_{j=1}^N$ と $Y$ の基底 $\{\mathbf{v}_k\}_{k=1}^M$ をとる．それぞれの双対基底を $\{\mathbf{f}^j\}_{j=1}^N$, $\{\mathbf{g}^k\}_{k=1}^M$ とおく．$\kappa_{jk}=\kappa(\mathbf{u}_j,\mathbf{v}_k)$ $(\in V)$ とおくと

$$\kappa_{jk}(\mathbf{f}^\mu,\mathbf{g}^\nu)=\begin{cases}1, & (j,k)=(\mu,\nu)\\ 0, & \text{その他}\end{cases}$$

である．まず $\{\kappa_{jk}\}$ は $V$ の基底になっていることを示す．$\varphi\in L(X^*,Y^*;\boldsymbol{K})$ とする．$c^{jk}=\varphi(\mathbf{f}^j,\mathbf{g}^k)$ とおくと，各 $\mathbf{f}^\mu$, $\mathbf{g}^\nu$ に対して

$$\sum_j\sum_k c^{jk}\kappa_{jk}(\mathbf{f}^\mu,\mathbf{g}^\nu)=c^{\mu\nu}=\varphi(\mathbf{f}^\mu,\mathbf{g}^\nu)$$

となっている．このことと $\sum_j\sum_k c^{jk}\kappa_{jk}$ および $\varphi$ の双線形性を用いて，$X^*\times Y^*$ 上で

$$\varphi=\sum_j\sum_k c^{jk}\kappa_{jk}$$

となることがわかる．またもしも $\sum_j\sum_k a^{jk}\kappa_{jk}=0$ ならば $(\mathbf{f}^\mu,\mathbf{g}^\nu)$ を代入して

$a^{\mu\nu}=0$ を得る．ゆえに $\{\kappa_{jk}\}$ は $V$ の基底である．このことから (1) は明らかである．

次に (2) を示す．$\boldsymbol{K}$ 上の任意の線形空間 $W$ と任意の $\Phi\in L(X,Y;W)$ をとる．$V\ni\varphi=\sum_j\sum_k c^{jk}\kappa_{jk}$ に対して，

$$T(\varphi)=\sum_j\sum_k c^{jk}\Phi(\mathbf{u}_j,\mathbf{v}_k)$$

と定義する．このとき，$T\in L(V;W)$ であり，$\mathbf{x}=\sum_j a^j\mathbf{u}_j$, $\mathbf{y}=\sum_k b^k\mathbf{v}_k$ に対して，

$$T(\kappa(\mathbf{x},\mathbf{y}))=\sum_j\sum_k a^j b^k\Phi(\mathbf{u}_j,\mathbf{v}_k)=\Phi(\mathbf{x},\mathbf{y})$$

となっている．

定理の後半の主張を証明する．$(V,\kappa)$, $(V',\kappa')$ が $X$ と $Y$ のテンソル積とする．$\kappa'\in L(X,Y;V')$ であるから，ある $T\in L(V;V')$ で $T\circ\kappa=\kappa'$ なるものが存在する．一方 $\kappa\in L(X,Y;V)$ であるから，ある $S\in L(V';V)$ が存在し，$S\circ\kappa'=\kappa$ をみたす．ゆえに $S\circ T\circ\kappa=\kappa$ が成り立つ．(1) の条件より $S\circ T=I_V$ ($I_V$ は $V$ 上の恒等写像 ) である．同様にして $T\circ S=I_{V'}$ も得られる．これより $T$ は $V$ から $V'$ への同形写像であることがわかる．$\chi=T$ とすればよい．■

テンソル積はただ一つとは限らないが, 定理 22.1 の後半の主張から, $X_1,\cdots,X_n$ の二つのテンソル積はテンソル積の意味で同形の違いを除いて一意的に定まっている．

**表記法 22. 4**　テンソル積 $(V,\kappa)$ について，慣用により $V$ を

$$X_1\otimes\cdots\otimes X_n$$

で表し，$\kappa(\mathbf{x}_1,\cdots,\mathbf{x}_n)$ を

$$\mathbf{x}_1\otimes\cdots\otimes\mathbf{x}_n$$

で表す．**ただしこの表し方はテンソル積の意味の同形に関して任意性がある**ので注意してほしい．

定理 22.1 の証明から次のことがわかる．

**定理 22. 5**　$X_1,\cdots,X_n$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．このとき次のこ

とが成り立つ.

(1)

$$X_1 \otimes \cdots \otimes X_n = L\left(X_1^*, \cdots, X_n^*; \boldsymbol{K}\right)$$

であり，かつ $\mathbf{x}_1 \in X_1, \cdots, \mathbf{x}_n \in X_n$, $\mathbf{g}^1 \in X_1^*, \cdots, \mathbf{g}^n \in X_n^*$ に対して，

$$\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n\left(\mathbf{g}^1, \cdots, \mathbf{g}^n\right) = \mathbf{g}^1\left(\mathbf{x}_1\right) \cdots \mathbf{g}^n\left(\mathbf{x}_n\right)$$

とみなすことができる.

(2) $\left\{\mathbf{u}_1^{(j)}, \cdots, \mathbf{u}_{N(j)}^{(j)}\right\}$ が $X_j\ (j = 1, \cdots, n)$ の基底ならば，

$$\mathbf{u}_{k(1)}^{(1)} \otimes \cdots \otimes \mathbf{u}_{k(n)}^{(n)}$$

$(k(j) = 1, \cdots, N(j),\ j = 1, \cdots, n)$ は $X_1 \otimes \cdots \otimes X_n$ の基底になっている．したがって

$$\dim X_1 \otimes \cdots \otimes X_n = \dim X_1 \times \cdots \times \dim X_n$$

である.

(3)

$$X_1 \otimes \cdots \otimes X_n = \operatorname{span}\left\{\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n : \mathbf{x}_1 \in X_1, \cdots, \mathbf{x}_n \in X_n\right\}$$

**注意 22.1** 例 22.1 で示した $\boldsymbol{K}^M$ と $\boldsymbol{K}^N$ のテンソル積 $\left(\boldsymbol{K}^{M \times N}, \kappa\right)$ に対して，$\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)} = \kappa\left(\mathbf{e}_j^{(M)}, \mathbf{e}_k^{(N)}\right)$ が $\boldsymbol{K}^{M \times N}$ の基底になっている．ここで $\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)}$ は $(j, k)$ 成分のみが 1 で，その他が 0 であるような $(M, N)$ 行列である.

**問題 22.1** $\left(\boldsymbol{K}^{M \times N}, \kappa\right)$ を例 22.1 で定めた $\boldsymbol{K}^M$ と $\boldsymbol{K}^N$ のテンソル積とする．$\mathbf{x} = \left(x_{jk}\right), \mathbf{y} = \left(y_{jk}\right) \in \boldsymbol{K}^{M \times N}$ に対して，

$$\langle \mathbf{x}, \mathbf{y} \rangle_{M \times N} = \sum_{j=1}^{M} \sum_{k=1}^{N} x_{jk} \overline{y_{jk}}$$

とすると，$\langle \cdot, \cdot \rangle_{M \times N}$ は $\boldsymbol{K}^{M \times N}$ の内積になっていることを示せ．また $\{\mathbf{u}_j\}_{j=1}^M$, $\{\mathbf{v}_k\}_{k=1}^N$ をそれぞれ $\boldsymbol{K}^M$, $\boldsymbol{K}^N$ の正規直交基底とするとき，$\{\mathbf{u}_j \otimes \mathbf{v}_k\}$ は内積 $\langle \cdot, \cdot \rangle_{M \times N}$ に関する $\boldsymbol{K}^{M \times N}$ の正規直交基底になっていることを示せ．ただし，$\boldsymbol{K}^M$, $\boldsymbol{K}^N$ の内積は定義 10.1 で定めた内積 $\langle \cdot, \cdot \rangle_{\boldsymbol{K}^M}$, $\langle \cdot, \cdot \rangle_{\boldsymbol{K}^N}$ とする.

三個以上の有限次元線形空間のテンソル積は逐次的に求めることもできる．簡単のため，三つの場合について述べる.

**定理 22.6** $X_1, X_2, X_3$ が $\boldsymbol{K}$ 上の有限次元線形空間であるとき，

$$\lambda(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3$$

により $\lambda$ を定義すると，$((X_1 \otimes X_2) \otimes X_3, \lambda)$ は $X_1, X_2, X_3$ のテンソル積になっている．したがって，

$$(X_1 \otimes X_2) \otimes X_3 \approx X_1 \otimes X_2 \otimes X_3.$$

**証明.** 明らかに $\lambda \in L(X_1, X_2, X_3; (X_1 \otimes X_2) \otimes X_3)$ である．テンソル積の定義のうち (1) が成り立つことは容易に示せる．

(2) を示す．$X_1 \otimes X_2 \otimes X_3$ は $X_1, X_2, X_3$ のテンソル積であるから，その定義より，ある $T \in L(X_1 \otimes X_2 \otimes X_3; (X_1 \otimes X_2) \otimes X_3)$ で

$$T(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3) = \lambda(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3$$

をみたすものが存在する．はじめに $T$ が同形写像であることを証明する．$\mathbf{x} \in X_3$ に対して，

$$\kappa_{\mathbf{x}}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}$$

とすると，$\kappa_{\mathbf{x}} \in L(X_1, X_2; X_1 \otimes X_2 \otimes X_3)$ であるから，ある $S_{\mathbf{x}} \in L(X_1 \otimes X_2; X_1 \otimes X_2 \otimes X_3)$ で

$$S_{\mathbf{x}}(\mathbf{x}_1 \otimes \mathbf{x}_2) = \kappa_{\mathbf{x}}(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}$$

をみたすものが存在する．いま

$$\Psi(\mathbf{y}, \mathbf{x}) = S_{\mathbf{x}}(\mathbf{y}), \quad \mathbf{y} \in X_1 \otimes X_2,\ \mathbf{x} \in X_3$$

とすると，$\Psi \in L((X_1 \otimes X_2), X_3; X_1 \otimes X_2 \otimes X_3)$ である．ゆえにある $S \in L((X_1 \otimes X_2) \otimes X_3; X_1 \otimes X_2 \otimes X_3)$ が存在し，

$$S((\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3) = \Psi(\mathbf{x}_1 \otimes \mathbf{x}_2, \mathbf{x}_3) = \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3.$$

ゆえに $S \circ T = I_{X_1 \otimes X_2 \otimes X_3}$ である．また $T \circ S = I_{(X_1 \otimes X_2) \otimes X_3}$ である．よって $T$ は $X_1 \otimes X_2 \otimes X_3$ から $(X_1 \otimes X_2) \otimes X_3$ への同形写像になっている．

$\Phi \in L(X_1, X_2, X_3; W)$ とする．このときある線形写像 $F \in L(X_1 \otimes X_2 \otimes X_3; W)$ が存在し，

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = F(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3)$$

である．さらに $F \circ S \in L((X_1 \otimes X_2) \otimes X_3; W)$ かつ

$$\begin{aligned}\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) &= F(\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3) \\ &= F \circ S((\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3)\end{aligned}$$

をみたしている．よって $(X_1 \otimes X_2) \otimes X_3, \lambda$ はテンソル積の定義 (2) もみたす．

■

この定理と同様にして，

$$\lambda'(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathbf{x}_1 \otimes (\mathbf{x}_2 \otimes \mathbf{x}_3)$$

により $\lambda'$ を定義すると，$(X_1 \otimes (X_2 \otimes X_3), \lambda')$ は $X_1, X_2, X_3$ のテンソル積になっている．したがって，

$$X_1 \otimes (X_2 \otimes X_3) \approx X_1 \otimes X_2 \otimes X_3$$

である．

**規約 22.7**　以下では $X_1 \otimes X_2 \otimes X_3$, $(X_1 \otimes X_2) \otimes X_3$, $X_1 \otimes (X_2 \otimes X_3)$ を同一視し，また

$$\mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3 = (\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3 = \mathbf{x}_1 \otimes (\mathbf{x}_2 \otimes \mathbf{x}_3)$$

とみなす (**テンソル積の結合律**).

## 22.2　線形写像のテンソル積

線形写像のテンソル積を定義する．まず次の定理を証明する．

**定理 22.8**　$X_1, Y_1, \cdots, X_n, Y_n$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．
$T_1 \in L(X_1; Y_1), \cdots, T_n \in L(X_n; Y_n)$ に対して，ある $T \in L(X_1 \otimes \cdots \otimes X_n; Y_1 \otimes \cdots \otimes Y_n)$ で

$$T(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n) = T_1(\mathbf{x}_1) \otimes \cdots \otimes T_n(\mathbf{x}_n), \tag{22.3}$$

$(\mathbf{x}_j \in X_j,\ j = 1, \cdots, n)$ をみたすものがただ一つ存在する．

**証明.** $n = 2$ の場合を証明する．一般の場合も同様にして示すことができる．$T_1 \in L(X_1; Y_1)$, $T_2 \in L(X_2; Y_2)$ に対して，まず

$$T_1 \times T_2(\mathbf{x}_1, \mathbf{x}_2) = T_1(\mathbf{x}_1) \otimes T_2(\mathbf{x}_2)$$

とすると，$T_1 \times T_2 \in L(X_1, X_2; Y_1 \otimes Y_2)$ である．このとき $X_1$ と $X_2$ のテンソル積の定義 (定義 22.1 (2) 参照) と補題 22.2 より，ある $T \in L(X_1 \otimes X_2; Y_1 \otimes Y_2)$ で，

$$T_1 \times T_2 (\mathbf{x}_1, \mathbf{x}_2) = T (\mathbf{x}_1 \otimes \mathbf{x}_2)$$

をみたすものがただ一つ存在する．この $T$ が求めるものである．もし $S, T \in L (X_1 \otimes X_2; Y_1 \otimes Y_2)$ が (22.3) をみたすならば，定理 22.5 (3) と写像の線形性から $S = T$ であることもわかる．■

**定義 22.3**　定理 22.8 の $T$ を $T_1, \cdots, T_n$ の**テンソル積**といい，

$$T = T_1 \otimes \cdots \otimes T_n$$

で表す．

じつは定義 22.3 で定められた $T_1 \otimes \cdots \otimes T_n$ により線形空間 $L (X_1; Y_1), \cdots,$ $L (X_n; Y_n)$ のテンソル積

$$L (X_1; Y_1) \otimes \cdots \otimes L (X_n; Y_n)$$

を定めることができる (問題 22.6 参照)．そのため，線形写像のテンソル積という言葉が用いられる．

**例 22.2**　$X$ を $\boldsymbol{K}$ 上の有限次元線形空間とし，$\mathbf{g}^1, \cdots, \mathbf{g}^n \in L(X; \boldsymbol{K}) = X^*$ とする．$T^n (X) = X \otimes \cdots \otimes X$ ($n$ 個のテンソル積) とする．このとき, $\mathbf{g}^1 \otimes \cdots \otimes \mathbf{g}^n \in L(T^n(X); \boldsymbol{K})$ であり，$\mathbf{x}_1, \cdots, \mathbf{x}_n \in X$ に対して

$$\mathbf{g}^1 \otimes \cdots \otimes \mathbf{g}^n (\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n) = \mathbf{g}^1 (\mathbf{x}_1) \cdots \mathbf{g}^n (\mathbf{x}_n)$$

である．また，$\mathbf{x}_j = \tau_{\mathbf{x}_j} \in L (X^*; \boldsymbol{K}) = (X^*)^*$ であり，$\tau_{\mathbf{x}_1} \otimes \cdots \otimes \tau_{\mathbf{x}_n} \in L (T^n (X^*); \boldsymbol{K})$ である．$X$ と $(X^*)^*$ は同一視して，$\mathbf{x}_j = \tau_{\mathbf{x}_j}$ とみなしているから，

$$\begin{aligned}\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n \left(\mathbf{g}^1 \otimes \cdots \otimes \mathbf{g}^n\right) &= \tau_{\mathbf{x}_1} \otimes \cdots \otimes \tau_{\mathbf{x}_n} \left(\mathbf{g}^1 \otimes \cdots \otimes \mathbf{g}^n\right) \\ &= \mathbf{g}^1 (\mathbf{x}_1) \cdots \mathbf{g}^n (\mathbf{x}_n) \\ &= \mathbf{g}^1 \otimes \cdots \otimes \mathbf{g}^n (\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_n)\end{aligned}$$

が成り立つ．

**例 22.3**　数ベクトル空間上で定義された線形写像のテンソル積の具体的な形を示しておこう．$T_A \in L \left(\boldsymbol{K}^M; \boldsymbol{K}^M\right)$ とし，$T_B \in L \left(\boldsymbol{K}^N; \boldsymbol{K}^N\right)$ とする．$\boldsymbol{K}^M \otimes \boldsymbol{K}^N$ として行列のなす線形空間 $\boldsymbol{K}^{M \times N}$ を考える．

このとき，$T_A \otimes T_B \in L \left(\boldsymbol{K}^{M \times N}; \boldsymbol{K}^{M \times N}\right)$ であり，

$$T_A \otimes T_B\,(\mathbf{x} \otimes \mathbf{y}) = T_A\,(\mathbf{x}) \otimes T_B\,(\mathbf{y})$$

をみたしている．特に $T_A$ の表現行列を $A = (\mathbf{a}_1, \cdots, \mathbf{a}_M)$ とすると，

$$T_A\left(\mathbf{e}_j^{(M)}\right) = A\mathbf{e}_j^{(M)} = \mathbf{a}_j$$

である．また $T_B$ の表現行列を $B = (\mathbf{b}_1, \cdots, \mathbf{b}_N)$ とすると，

$$T_B\left(\mathbf{e}_k^{(N)}\right) = B\mathbf{e}_k^{(N)} = \mathbf{b}_k$$

である．したがって，

$$T_A \otimes T_B\left(\mathbf{e}_j^{(M)} \otimes \mathbf{e}_j^{(M)}\right) = \mathbf{a}_j \otimes \mathbf{b}_k$$

となっている．一般に $\boldsymbol{K}^{M\times N} \ni \mathbf{x} = \sum_{j=1}^{M}\sum_{k=1}^{N} x_{jk}\mathbf{e}_j^{(M)} \otimes \mathbf{e}_j^{(M)}$ と表せるから，

$$T_A \otimes T_B\,(\mathbf{x}) = \sum_{j=1}^{M}\sum_{k=1}^{N} x_{jk}\mathbf{a}_j \otimes \mathbf{b}_k$$

と表せることがわかる．

ところで，$(a_{jk}) \in \boldsymbol{K}^{M\times N}$ を

$$(a_{ij}) = (a_{11}, \cdots, a_{1N}, a_{21}, \cdots, a_{2N}, \cdots, a_{M1}, \cdots, a_{MN})^T$$

と並べることにより，$\boldsymbol{K}^{M\times N}$ を $\boldsymbol{K}^{MN}$ とみなすことができる．この並べ方を**辞書的配列**という．辞書的配列により $T_A \otimes T_B \in L\left(\boldsymbol{K}^{MN}; \boldsymbol{K}^{MN}\right)$ とみなすこともできる．したがって $C \in \boldsymbol{K}^{MN\times MN}$ で，

$$T_A \otimes T_B = T_C$$

をみたすものが存在する．それはどのような行列だろうか．これは練習問題とする (問題 22.7).

線形写像のテンソル積の合成について次のことが成り立つ.

**定理 22.9** $X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．$T_1, T_2 \in L\,(X; X)$ , $S_1, S_2 \in L\,(Y; Y)$ とする．

$$A_1 = T_1 \otimes S_1, \quad A_2 = T_2 \otimes S_2 \in L\,(X \otimes Y; X \otimes Y)$$

とする．このとき

$$A_2 \circ A_1 = (T_2 \circ T_1) \otimes (S_2 \circ S_1)$$

となる．

**証明.** $\mathbf{z} = \mathbf{x} \otimes \mathbf{y} \in X \otimes Y$ に対して,

$$\begin{aligned} A_2 \circ A_1 (\mathbf{z}) &= A_2 (T_1 (\mathbf{x}) \otimes S_1 (\mathbf{y})) = T_2 (T_1 (\mathbf{x})) \otimes S_2 (S_1 (\mathbf{y})) \\ &= (T_2 \circ T_1) \otimes (S_2 \circ S_1) (\mathbf{z}) \end{aligned}$$

である. $X \otimes Y$ のすべての要素は $\mathbf{x} \otimes \mathbf{y}$ の形の要素の一次結合で表されているので, 写像の線形性から求める等式が得られる. ■

この定理から次の定理を証明することができる.

**定理 22.10**　$X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする. $T \in L(X; X)$, $S \in L(Y; Y)$ とし, 逆写像 $T^{-1}$, $S^{-1}$ が存在するものとする. このとき $T \otimes S \in L(X \otimes Y; X \otimes Y)$ の逆写像が存在し,

$$(T \otimes S)^{-1} = T^{-1} \otimes S^{-1}$$

で与えられる.

**証明.** $R = T^{-1} \otimes S^{-1}$ とすると, $R \in L(X \otimes Y; X \otimes Y)$ であり, 定理 22.9 より

$$R \circ (T \otimes S) = \left(T^{-1} \circ T\right) \otimes \left(S^{-1} \circ S\right) = Id_X \otimes Id_Y = Id_{X \otimes Y}.$$

同様にして $(T \otimes S) \circ R = Id_{X \otimes Y}$ も成り立つ. ■

**定理 22.11（逐次計算法）**　$X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする. $T \in L(X; X)$, $S \in L(Y; Y)$ とする. このとき, $\mathbf{x} \in X \otimes Y$ に対して

$$\begin{aligned} T \otimes S (\mathbf{x}) &= T \otimes Id_Y \ (Id_X \otimes S (\mathbf{x})) \\ &= Id_X \otimes S \ (T \otimes Id_Y (\mathbf{x})) \end{aligned}$$

が成り立つ ($Id_X, Id_Y$ はそれぞれ $X, Y$ の恒等写像を表している ).

具体的な場合に定理 22.2 を用いて, 線形写像のテンソル積の計算方法を述べておこう. 以下の計算方法は, 離散フーリエ変換やウェーブレット変換のテンソル積を計算する際に用いられる. これらの変換のテンソル積は画像解析で本質的な役割を果たす.

$\mathbf{x} = (x_{jk}) \in \boldsymbol{K}^{M \times N}$ とする. まず $\mathbf{x}$ の第 $j$ 行 $x(j, :)$ をたてに並べ替えて, $N$ 次元数ベクトルとみなす. これに $T_B$ を施し (この操作を $\mathbf{x}$ の第 $j$ 行に $T_B$ を作用させるという),

$$\begin{pmatrix} y_{j1} \\ \vdots \\ y_{jN} \end{pmatrix} = T_B \begin{pmatrix} x_{j1} \\ \vdots \\ x_{jN} \end{pmatrix}$$

とおく．これを各行に対して行う．ここで得られたベクトルを転置してから並べて，次のような行列を定義する．

$$\mathbf{y} = \begin{pmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{M1} & \cdots & y_{MN} \end{pmatrix} \in \boldsymbol{K}^{M\times N}.$$

次に $\mathbf{y}$ の各 $k$ 列 $y(:,k)$ を $M$ 次元数ベクトルとみなして $T_A$ を施す (この操作を $\mathbf{y}$ の第 $k$ 列に $T_A$ を作用させるという)．その結果を

$$\begin{pmatrix} z_{1k} \\ \vdots \\ z_{Mk} \end{pmatrix} = T_A \begin{pmatrix} y_{1k} \\ \vdots \\ y_{Mk} \end{pmatrix} \tag{22.4}$$

とおく．次の定理が成り立つ．

**定理 22. 12**　(22.4) で定義した $(z_{jk})$ に対して，

$$T_A \otimes T_B\,(\mathbf{x}) = \begin{pmatrix} z_{11} & \cdots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{M1} & \cdots & z_{MN} \end{pmatrix}.$$

**証明.** $I_M$ は $M$ 次単位行列あるいは $\boldsymbol{K}^M$ の恒等写像を表すものとする．

$$\begin{pmatrix} & O & \\ x_{j1} & \cdots & x_{jN} \\ & O & \end{pmatrix} = \sum_{k=1}^{N} x_{jk}\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)}$$

である．ゆえに

$$I_M \otimes T_B \left(\sum_{k=1}^{N} x_{jk}\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)}\right) = \mathbf{e}_j^{(M)} \otimes T_B \left(\sum_{k=1}^{N} x_{jk}\mathbf{e}_k^{(N)}\right)$$

$$= \mathbf{e}_j^{(M)} \otimes \sum_{k=1}^{N} y_{jk}\mathbf{e}_k^{(N)} = \begin{pmatrix} & O & \\ y_{j1} & \cdots & y_{jN} \\ & O & \end{pmatrix}.$$

また

$$
\begin{aligned}
I_M \otimes T_B(\mathbf{x}) &= I_M \otimes T_B\left(\sum_{j=1}^{M}\sum_{k=1}^{N} x_{jk}\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)}\right) \\
&= \sum_{j=1}^{M} I_M \otimes T_B\left(\sum_{k=1}^{N} x_{jk}\mathbf{e}_j^{(M)} \otimes \mathbf{e}_k^{(N)}\right) \\
&= \begin{pmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{M1} & \cdots & y_{MN} \end{pmatrix}.
\end{aligned}
$$

したがって，

$$
\begin{aligned}
T_A \otimes I_N\left(I_M \otimes T_B(\mathbf{x})\right) &= \sum_{k=1}^{N} T_A \otimes I_N\left(\left(\sum_{j=1}^{N} y_{jk}\mathbf{e}_j^{(M)}\right) \otimes \mathbf{e}_k^{(N)}\right) \\
&= \sum_{k=1}^{N}\left(\sum_{j=1}^{M} z_{jk}\mathbf{e}_j^{(M)}\right) \otimes \mathbf{e}_k^{(N)} \\
&= \begin{pmatrix} z_{11} & \cdots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{M1} & \cdots & z_{MN} \end{pmatrix}.
\end{aligned}
$$

■

**問題 22. 2**　$\mathbf{x}, \mathbf{x}' \in X$, $\mathbf{y}, \mathbf{y}' \in Y$, $\alpha, \alpha' \in \boldsymbol{K}$ に対して

$$
\begin{aligned}
(\alpha\mathbf{x} + \alpha'\mathbf{x}') \otimes \mathbf{y} &= \alpha(\mathbf{x} \otimes \mathbf{y}) + \alpha'(\mathbf{x}' \otimes \mathbf{y}), \\
\mathbf{x} \otimes (\alpha\mathbf{y} + \alpha'\mathbf{y}') &= \alpha(\mathbf{x} \otimes \mathbf{y}) + \alpha'(\mathbf{x} \otimes \mathbf{y}')
\end{aligned}
$$

を示せ (**テンソル積の双線形性**).

**問題 22. 3**　$X, Y, V$ を $\boldsymbol{K}$ 上の有限次元線形空間で，$M = \dim X$, $N = \dim Y$, $MN = \dim V$ とする．

$\{\mathbf{u}_j\}_{j=1}^{M}$, $\{\mathbf{v}_k\}_{k=1}^{N}$, $\{\mathbf{e}_{jk}\}_{j=1,\cdots M,\, k=1,\cdots,N}$ をそれぞれ $X, Y, V$ の基底とする．$\mathbf{x} = \sum_{j=1}^{M} \alpha^j \mathbf{u}_j$, $\mathbf{y} = \sum_{k=1}^{N} \beta^k \mathbf{v}_k$ に対して，

$$
\kappa(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{M}\sum_{k=1}^{N} \alpha^j \beta^k \mathbf{e}_{jk}
$$

とする．このとき $(V, \kappa)$ は $X$ と $Y$ のテンソル積である．

**問題 22.4** $X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．$\mathbf{f} \in X^*$, $\mathbf{y} \in Y$ に対して

$$\tau_{\mathbf{y},\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})\,\mathbf{y}, \quad \mathbf{x} \in X$$

とし，$\tau(\mathbf{y}, \mathbf{f}) = \tau_{\mathbf{y},\mathbf{f}}$ とする．このとき $(L(X;Y), \tau)$ は $Y$ と $X^*$ のテンソル積になっていることを示せ．したがって

$$L(X;Y) \approx Y \otimes X^*$$

である．

**問題 22.5** $X, Y$ を $\boldsymbol{K}$ 上の有限次元線形空間とする．このとき $\dim L(X;Y) = \dim X \cdot \dim Y$ であり，$Y$ の基底 $\mathbf{y}_1, \cdots, \mathbf{y}_N$，$X^*$ の基底 $\mathbf{f}^1, \cdots, \mathbf{f}^M$ に対して，

$$\mathbf{y}_k \otimes \mathbf{f}^j(\mathbf{x}) = \mathbf{f}^j(\mathbf{x})\mathbf{y}_k, \quad \mathbf{x} \in X$$

として定義される $\mathbf{y}_k \otimes \mathbf{f}^j$ が $L(X;Y)$ の基底になっている．

**問題 22.6** $T \in L(X_1;Y_1)$, $S \in L(X_2;Y_2)$ に対して，

$$\kappa(T, S) = T \otimes S$$

とする．このとき，

$$\kappa \in L\left(L(X_1;Y_1), L(X_2;Y_2); L(X_1 \otimes X_2; Y_1 \otimes Y_2)\right)$$

であり，さらに $(L(X_1 \otimes X_2, Y_1 \otimes Y_2), \kappa)$ は $L(X_1;Y_1)$ と $L(X_2;Y_2)$ のテンソル積になっている．

**問題 22.7**

$$A = \left(a^j_k\right)_{j=1,\cdots,N_1,\, k=1,\cdots,M_1} \in \boldsymbol{K}^{N_1 \times M_1},$$
$$B = \left(b^j_k\right)_{j=1,\cdots,N_2,\, k=1,\cdots,M_2} \in \boldsymbol{K}^{N_2 \times M_2}$$

とし,それから定まる線形写像をそれぞれ $T_A \in L(\boldsymbol{K}^{M_1};\boldsymbol{K}^{N_1})$, $T_B \in L(\boldsymbol{K}^{M_2};\boldsymbol{K}^{N_2})$ とする．このとき

$$T_A \otimes T_B \in L(\boldsymbol{K}^{M_1} \otimes \boldsymbol{K}^{M_2}; \boldsymbol{K}^{N_1} \otimes \boldsymbol{K}^{N_2})$$

である．また $\boldsymbol{K}^{M_1 \times N_1}$ の要素を辞書的配列することにより $\boldsymbol{K}^{M_1 \times N_1}$ を $\boldsymbol{K}^{M_1 N_1}$ とみなす．同様に $\boldsymbol{K}^{N_1} \otimes \boldsymbol{K}^{N_2}$ を $\boldsymbol{K}^{N_1 N_2}$ とみなす．このとき，$T_A \otimes T_B \in L(\boldsymbol{K}^{M_1 M_2}; \boldsymbol{K}^{N_1 N_2})$ となっている．したがって，ある $C \in \boldsymbol{K}^{N_1 N_2 \times M_1 M_2}$ により，

$$T_A \otimes T_B = T_C$$

となる．この行列 $C$ は

$$C = \begin{pmatrix} a_1^1 B & a_2^1 B & \cdots & a_{M_1}^1 B \\ a_1^2 B & a_2^2 B & \cdots & a_{M_1}^2 B \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{N_1} B & a_2^{N_1} B & \cdots & a_{M_1}^{N_1} B \end{pmatrix}$$

であることを示せ．$C$ を $A$ と $B$ の**クロネッカー積**という．

## 22.3　画像処理と線形写像のテンソル積

画像は数の 2 次元的な配列として表される．たとえば $M \times N$ ピクセルのグレイ・スケールの画像 $\mathbf{x}$ は，

$$\mathbf{x} = \begin{pmatrix} x[0,0] & x[0,1] & \cdots & x[0,N-1] \\ x[1,0] & x[1,1] & \cdots & x[1,N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x[M-1,0] & x[M-1,1] & \cdots & x[M-1,N-1] \end{pmatrix}$$

と表される．ただし $x[j,k]$ は画像の位置 $(j,k)$ における濃淡を表す数値である．たとえば図 1.2 の場合，画像は $256 \times 256$ ピクセル，256 階調のディジタル画像で，$x[j,k] \in \{0,1,2,\cdots,255\}$ となっていて，0 は黒を表し，255 は白を表す．0 から 255 に整数が大きくなるにつれ，次第に薄い灰色を表す．実際に図 1.2 は

$$x = \begin{pmatrix} 6 & 12 & 12 & \cdots & 14 & 12 & 16 \\ 7 & 2 & 1 & \cdots & 2 & 0 & 4 \\ 5 & 1 & 1 & \cdots & 4 & 1 & 5 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 128 & 100 & 86 & \cdots & 111 & 106 & 86 \\ 104 & 90 & 85 & \cdots & 121 & 112 & 95 \\ 104 & 105 & 88 & \cdots & 122 & 118 & 96 \end{pmatrix}$$

のように表されている．したがって，$M \times N$ ピクセルのグレイ・スケールの画像は $\boldsymbol{K}^{M \times N}$ に属する行列とみなすことができる．このような画像にフーリエ変換，ウェーブレット変換などを施したい場合，$\boldsymbol{K}^{M \times N}$ をテンソル積 $\boldsymbol{K}^M \otimes \boldsymbol{K}^N$ とみなすと便利である．たとえば $M$ 点離散フーリエ変換 $\mathcal{F}_M$ と $N$ 点離散フーリエ変換 $\mathcal{F}_N$ はそれぞれ $\mathcal{F}_M \in L\left(\boldsymbol{K}^M; \boldsymbol{K}^M\right)$, $\mathcal{F}_N \in L\left(\boldsymbol{K}^N; \boldsymbol{K}^N\right)$ であるから，

離散フーリエ変換のテンソル積は

$$\mathcal{F}_M \otimes \mathcal{F}_N \in L\left(\boldsymbol{K}^M \otimes \boldsymbol{K}^N; \boldsymbol{K}^M \otimes \boldsymbol{K}^N\right)$$

となっている．$\mathcal{F}_{M,N} = \mathcal{F}_M \otimes \mathcal{F}_N$ と表す．$\mathcal{F}_{M,N}$ を画像 $\mathbf{x}$ に施すことができる．

**注意 22.2**　ここでの設定であると，データは周期拡張されているが，画像データの解析では，周期拡張すると境界での処理に不自然さがある．そのため，実際には対称拡張などをすることが多い．

最後に，離散フーリエ変換を用いた画像処理例を述べておく．定理 22.10 より

$$\mathcal{F}_{M,N}^{-1} = \mathcal{F}_M^{-1} \otimes \mathcal{F}_N^{-1}$$

は $\mathcal{F}_{M,N}$ の逆写像になっている．いま $\mathbf{m} \in \boldsymbol{K}^M \otimes \boldsymbol{K}^N = \boldsymbol{K}^{M\times N}$ とし，

$$T_{\mathbf{m}}(\mathbf{x}) = \mathcal{F}_{M,N}^{-1}\left(\mathbf{m}_\bullet \mathcal{F}_{M,N}(\mathbf{x})\right), \quad \mathbf{x} \in \boldsymbol{K}^{M\times N}$$

と定義する．これは第 18.3 節の定義 18.5 で定義したフーリエ・マルチプライヤのテンソル積版である．命題 18.3 の後で述べたように，$\mathbf{f} \in \boldsymbol{K}^M$ の $M$ 点離散フーリエ変換 $\mathcal{F}_M(\mathbf{f})[k]$ は，主に $k$ が $\dfrac{M}{2}$ の周辺にある場合は $\mathbf{f}$ の高周波部分を表し，$k$ が 0 あるいは $M-1$ のような端点の周辺にある場合は低周波部分を表す．したがってテンソル積 $\mathcal{F}_{M,N}$ の場合，$\mathcal{F}_{M,N}(\mathbf{x})[j,k]$ は $j$ と $k$ がそれぞれ $\dfrac{M}{2}$, $\dfrac{N}{2}$ の周辺にある場合が $x$ の高周波部分，そして，$j, k$ が端点の周辺にある場合が低周波部分を表している．画像における高周波部分とは，大雑把にいえば，画像の模様が細かい部分で，低周波部分とは模様がそれほど細かくない部分である．

$\delta_1 > 0, \delta_2 > 0$ を小さい自然数とし，

$$\begin{aligned}
E_{00} &= \{(j,k) : |j| < \delta_1,\ |k| < \delta_1\} \\
E_{10} &= \{(j,k) : |j-(M-1)| < \delta_1,\ |k| < \delta_1\} \\
E_{01} &= \{(j,k) : |j| < \delta_1,\ |k-(N-1)| < \delta_1\} \\
E_{11} &= \{(j,k) : |j-(M-1)| < \delta_1,\ |k-(N-1)| < \delta_1\} \\
C &= \left\{(j,k) : \left|j-\frac{M}{2}\right| < \delta_2,\ \left|k-\frac{N}{2}\right| < \delta_2\right\}
\end{aligned}$$

とする．$\mathbf{m} = (m_{jk}) \in \boldsymbol{K}^{M\times N}$ として，たとえば $(j,k) \in C$ のときは $|m_{jk}|$ は大きい正の値をとるようにしておくと，$T_{\mathbf{m}}(\mathbf{x})$ は $\mathbf{x}$ の高周波部分を強調する操作を行う．たとえば，いま $\mathbf{x}$ として図 22.1 の左図の画像とする．この画像は $500\times 500$ 画素の画像である．

次に



図 22.1　左図：原画像，右図：高周波強調画像

$$\mathbf{a} = \begin{pmatrix} 0 & -1 & 0 & \\ -1 & 4 & -1 & O \\ 0 & -1 & 0 & \\ & & & \\ & O & & O \\ & & & \end{pmatrix} \in \boldsymbol{K}^{500\times 500}$$

とし，$\mathbf{m} = \mathcal{F}_{500,500}(\mathbf{a})$ とする．このとき，$(|\mathcal{F}_{500,500}(\mathbf{a})[j,k]|) \in \boldsymbol{K}^{500\times 500}$ は図 22.2 のようなグラフをもつ．

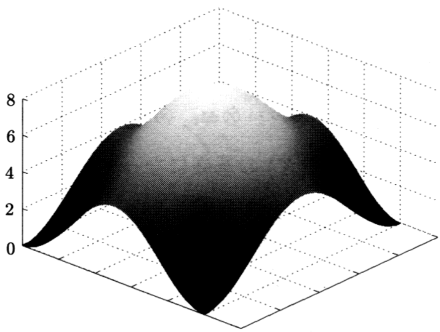



図 22.2

図 22.2 からわかるように $\mathbf{m}$ は高周波を抜き出す作用がある．そこで，

$$\mathbf{x} + T_{\mathbf{m}}(\mathbf{x})$$

の実数部分を $\mathbf{y}$ とおくと，これは $\mathbf{x}$ の高周波を強調した画像になっている．図22.1 の右図が $\mathbf{y}$ を表している．ただし，$y$ の成分には負数も含まれているため，スケーリングをして表示している．

この操作により画像の細かい見にくい部分をより鮮鋭にすることができる．

このほか，たとえば高周波部分をカットすることにより画像圧縮も可能である．JPEG では離散コサイン変換 (例 1.6) のテンソル積が用いられている．

## 22.4　反変テンソル，共変テンソル

本節を通して，$X$ を $\boldsymbol{K}$ 上の $N$ 次元線形空間とする．$X$ の $p$ 個のテンソル積と $X^*$ の $q$ 個のテンソル積のテンソル積を

$$T_q^p(X) = \underbrace{X \otimes \cdots \otimes X}_{p\ 個} \otimes \underbrace{X^* \otimes \cdots \otimes X^*}_{q\ 個}$$

と表す．特に $X$ の $p$ 個のテンソル積は $T^p(X)$ または $T_0^p(X)$ と表す．同様にして $T_q(X) = T_q^0(X)$ は $X^*$ の $q$ 個のテンソル積とする．$T_q^p(X)$ に属する要素を **$p$ 階反変 $q$ 階共変テンソル**という．また $T^p(X)$ の要素は **$p$ 階反変テンソル**，$T_q(X)$ を **$q$ 階共変テンソル**という．また

$$T^0(X) = T_0(X) = T_0^0(X) = \boldsymbol{K}$$

とする．

テンソルの基底変換に対する変換則を計算しておこう．$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とし，$\left\{\mathbf{f}^1, \cdots, \mathbf{f}^N\right\}$ をその双対基底とする．このとき，

$$\mathbf{u}_{i_1 \cdots i_p}^{j_1 \cdots j_q} = \mathbf{u}_{i_1} \otimes \cdots \otimes \mathbf{u}_{i_p} \otimes \mathbf{f}^{j_1} \otimes \cdots \otimes \mathbf{f}^{j_q} \tag{22.5}$$

$(i_1, \cdots, j_q \in \{1, \cdots, N\})$ は $T_q^p(X)$ の基底になっている．これを $\mathbf{u}_1, \cdots, \mathbf{u}_N$ から定義される $T_q^p(X)$ の**基底**という．$\mathbf{t} \in T_q^p(X)$ は

$$\mathbf{t} = \sum_{i_1, \cdots, i_p, j_1, \cdots, j_q = 1}^{N} t_{j_1 \cdots j_q}^{i_1 \cdots i_p} \, \mathbf{u}_{i_1 \cdots i_p}^{j_1 \cdots j_q}$$

と表せる．いま $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ を $X$ の基底とする．このとき $\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N$ から同様の手順で定義される $T_q^p(X)$ の基底を $\widetilde{\mathbf{u}}_{i_1 \cdots i_p}^{j_1 \cdots j_q}$ とすると，

$$\mathbf{t} = \sum_{i_1, \cdots, i_p, j_1, \cdots, j_q = 1}^{N} \widetilde{t}_{j_1 \cdots j_q}^{i_1 \cdots i_p} \, \widetilde{\mathbf{u}}_{i_1 \cdots i_p}^{j_1 \cdots j_q}$$

と表せる．本節ではこの二つの係数 $t^{i_1\cdots i_p}_{j_1\cdots j_q}$ と $\widetilde{t}^{i_1\cdots i_p}_{j_1\cdots j_q}$ の間の関係式を求める．基底の間に

$$\widetilde{\mathbf{u}}_i = \sum_{\nu=1}^{N} u_i^{\nu}\mathbf{u}_{\nu} \tag{22.6}$$

の関係式が成り立っているとする．$A = \left(u_j^i\right)$ を $(i,j)$ 成分が $u_j^i$ の $N$ 次正方行列とする．このとき $A$ は正則行列であり，その逆行列を $A^{-1} = \left(\widetilde{u}_j^i\right)$ とする (第 21.3 節で述べたことを以下では用いる)．

**定理 22. 13**　この二つの係数 $t^{i_1\cdots i_p}_{j_1\cdots j_q}$ と $\widetilde{t}^{i_1\cdots i_p}_{j_1\cdots j_q}$ の間には

$$\widetilde{t}^{i_1\cdots i_p}_{j_1\cdots j_q} = \sum_{\nu_1,\cdots,\nu_p,\mu_1,\cdots,\mu_q=1}^{N} \widetilde{u}^{i_1}_{\nu_1}\cdots\widetilde{u}^{i_p}_{\nu_p}u^{\mu_1}_{j_1}\cdots u^{\mu_p}_{j_p}t^{\nu_1\cdots\nu_p}_{\mu_1\cdots\mu_q} \tag{22.7}$$

の関係式が成り立っている．

**証明.** (21.4), (21.7) を用いて次の式変形ができる．

$$\begin{aligned}
\mathbf{u}^{\mu_1\cdots\mu_q}_{\nu_1\cdots\nu_p} &= \mathbf{u}_{\nu_1}\otimes\cdots\otimes\mathbf{u}_{\nu_p}\otimes\mathbf{f}^{\mu_1}\otimes\cdots\otimes\mathbf{f}^{\mu_q}\\
&= \sum_{i_1,\cdots,i_p,j_1,\cdots,j_q=1}^{N} \widetilde{u}^{i_1}_{\nu_1}\cdots\widetilde{u}^{i_p}_{\nu_p}u^{\mu_1}_{j_1}\cdots u^{\mu_p}_{j_p}\widetilde{\mathbf{u}}_{i_1}\otimes\cdots\otimes\widetilde{\mathbf{u}}_{i_p}\otimes\widetilde{\mathbf{f}}^{j_1}\otimes\cdots\otimes\widetilde{\mathbf{f}}^{j_q}\\
&= \sum_{i_1,\cdots,i_p,j_1,\cdots,j_q=1}^{N} \widetilde{u}^{i_1}_{\nu_1}\cdots\widetilde{u}^{i_p}_{\nu_p}u^{\mu_1}_{j_1}\cdots u^{\mu_p}_{j_p}\widetilde{\mathbf{u}}^{j_1\cdots j_p}_{i_1\cdots i_p}.
\end{aligned}$$

ゆえに

$$\sum_{i_1,\cdots,i_p,j_1,\cdots,j_q=1}^{N} \widetilde{t}^{i_1\cdots i_p}_{j_1\cdots j_q}\,\widetilde{\mathbf{u}}^{j_1\cdots j_q}_{i_1\cdots i_p} = \mathbf{t} = \sum_{\nu_1,\cdots,\nu_p,\mu_1,\cdots,\mu_q=1}^{N} t^{\nu_1\cdots\nu_p}_{\mu_1\cdots\mu_q}\,\mathbf{u}^{\mu_1\cdots\mu_q}_{\nu_1\cdots\nu_p}$$

$$= \sum_{\nu_1,\cdots,\nu_p,\mu_1,\cdots,\mu_q=1}^{N} t^{\nu_1\cdots\nu_p}_{\mu_1\cdots\mu_q} \sum_{i_1,\cdots,i_p,j_1,\cdots,j_q=1}^{N} \widetilde{u}^{i_1}_{\nu_1}\cdots\widetilde{u}^{i_p}_{\nu_p}u^{\mu_1}_{j_1}\cdots u^{\mu_p}_{j_p}\widetilde{\mathbf{u}}^{j_1\cdots j_p}_{i_1\cdots i_p}.$$

したがって

$$\widetilde{t}^{i_1\cdots i_p}_{j_1\cdots,j_q} = \sum_{\nu_1,\cdots,\nu_p,\mu_1,\cdots,\mu_q=1}^{N} \widetilde{u}^{i_1}_{\nu_1}\cdots\widetilde{u}^{i_p}_{\nu_p}u^{\mu_1}_{j_1}\cdots u^{\mu_p}_{j_p}t^{\nu_1\cdots\nu_p}_{\mu_1\cdots\mu_q}.$$

■

## 22.5 交代テンソル

本節を通して，$X$ を $\boldsymbol{K}$ 上の $N$ 次元線形空間とする．$S_p$ を $p$ 次置換の全体のなす集合とする．$\sigma \in S_p$ に対して，

$$P_\sigma(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_p) = \mathbf{x}_{\sigma^{-1}(1)} \otimes \cdots \otimes \mathbf{x}_{\sigma^{-1}(p)}$$

により，$T^p(X)$ から $T^p(X)$ への線形写像 $P_\sigma$ を定義できる．

**定義 22.4** $\mathbf{t} \in T^p(X)$ は，すべての置換 $\sigma \in S_p$ に対して

$$P_\sigma(\mathbf{t}) = \mathbf{t}$$

をみたすとき，$p$ **階対称テンソル**という．またすべての置換 $\sigma \in S_p$ に対して

$$P_\sigma(\mathbf{t}) = \mathrm{sgn}(\sigma)\, \mathbf{t}$$

をみたすとき,$p$ **階交代テンソル**という．$p$ 階対称テンソル全体のなす集合を $S^p(X)$，$p$ 階交代テンソル全体のなす集合を $\Lambda^p(X)$ と表す．

$\mathbf{t} \in T^p(X)$ に対して，

$$\mathcal{S}_p(\mathbf{t}) = \frac{1}{p!} \sum_{\sigma \in S_p} P_\sigma(\mathbf{t}),$$
$$\mathcal{A}_p(\mathbf{t}) = \frac{1}{p!} \sum_{\sigma \in S_p} \mathrm{sgn}(\sigma) P_\sigma(\mathbf{t})$$

と定める．このとき，$\mathcal{S}_p(\mathbf{t})$ は $p$ 階対称テンソル，$\mathcal{A}_p(\mathbf{t})$ は $p$ 階交代テンソルである．さらに次のことが容易に証明できる．

**定理 22.14** (1) $\mathbf{t} \in T^p(X)$ に対して，

$$\mathcal{S}_p(\mathcal{S}_p(\mathbf{t})) = \mathcal{S}_p(\mathbf{t}), \quad \mathcal{A}_p(\mathcal{A}_p(t)) = \mathcal{A}_p(\mathbf{t}).$$

(2)

$$S^p(X) = \{\mathcal{S}_p(\mathbf{t}) : \mathbf{t} \in T^p(X)\} = \{\mathbf{t} \in T^p(X) : \mathcal{S}_p(\mathbf{t}) = \mathbf{t}\},$$
$$\Lambda^p(X) = \{\mathcal{A}_p(\mathbf{t}) : \mathbf{t} \in T^p(X)\} = \{\mathbf{t} \in T^p(X) : \mathcal{A}_p(\mathbf{t}) = \mathbf{t}\}.$$

(3) $S^1(X) = \Lambda^1(X) = T^1(X) = X$ であり，$p \geq 2$ の場合は

$$S^p(X) \cap \Lambda^p(X) = \{\mathbf{0}\}.$$

$X$ の基底 $\mathbf{u}_1, \cdots, \mathbf{u}_N$ から定義される $T^p(X)$ の基底を $\mathbf{u}_{i_1 \cdots i_p}\,(i_1, \cdots, i_p \in \{1, \cdots, N\})$ とする．

$T^p(X) \ni \mathbf{t} = \sum_{i_1,\cdots,i_p=1}^{N} t^{i_1\cdots i_p}\mathbf{u}_{i_1\cdots i_p}$ とすると

$$P_\sigma(\mathbf{t}) = \sum_{i_1,\cdots,i_p=1}^{N} t^{i_1\cdots i_p}\mathbf{u}_{\sigma^{-1}(i_1)\cdots\sigma^{-1}(i_p)} = \sum_{i_1,\cdots,i_p=1}^{N} t^{\sigma(i_1)\cdots\sigma(i_p)}\mathbf{u}_{i_1\cdots i_p}$$

であるから次のことがわかる．

**定理 22.15**　$T^p(X) \ni \mathbf{t} = \sum_{i_1,\cdots,i_p=1}^{N} t^{i_1\cdots i_p}\mathbf{u}_{i_1\cdots i_p}$ とする．$\mathbf{t}$ が対称テンソルであるための必要十分条件は，任意の $\sigma \in S_p$ に対して，

$$t^{\sigma(i_1)\cdots\sigma(i_p)} = t^{i_1\cdots i_p}$$

となることである．また $\mathbf{t}$ が交代テンソルであるための必要十分条件は，任意の $\sigma \in S_p$ に対して，

$$t^{\sigma(i_1)\cdots\sigma(i_p)} = \mathrm{sgn}(\sigma)\, t^{i_1\cdots i_p}$$

となることである．

本節の最後に $S^p(X)$ および $\Lambda^p(X)$ の次元を求めておこう．

**定理 22.16**　(1)　$p \geq 1$ に対して，

$$\dim S^p(X) = \binom{N+p-1}{p} = \frac{(N+p-1)!}{(N-1)!\,p'}.$$

(2)　$1 \leq p \leq N$ の場合，

$$\dim \Lambda^p(X) = \binom{N}{p} = \frac{N!}{(N-p)!\,p!}.$$

また $p > N$ の場合は $\dim \Lambda^p(X) = 0$ である．

**証明．** (1)　$S^p(X)$ は $\{\mathcal{S}_p(\mathbf{u}_{i_1\cdots i_p}) : i_1, \cdots, i_p \in \{1, \cdots, N\}\}$ で生成されている．いま重複度も込めて集合として $\{i_1, \cdots, i_p\} = \{j_1, \cdots, j_p\}$ であるとき，ある $\sigma \in S_p$ で $i_{\sigma(k)} = j_k$, $k = 1, \cdots, p$ となるものが存在する．

$$\mathcal{S}_p(\mathbf{u}_{i_1\cdots i_p}) = \frac{1}{p!}\sum_{\tau \in S_p} \mathbf{u}_{\tau(i_1)\cdots\tau(i_p)}$$

であるから，$\mathcal{S}_p(\mathbf{u}_{i_1\cdots i_p}) = \mathcal{S}_p(\mathbf{u}_{j_1\cdots j_p})$ である．一方，$\mathcal{S}_p(\mathbf{u}_{i_1\cdots i_p}) = \mathcal{S}_p(\mathbf{u}_{j_1\cdots j_p})$ ならば，$i_1, \cdots, i_p$ と $j_1, \cdots, j_p$ は順番を入れ替えたものでなければならないので，重複度も込めて集合として $\{i_1, \cdots, i_p\} = \{j_1, \cdots, j_p\}$ である．また，重複度も込めて集合として $\{i_1, \cdots, i_p\} \neq \{j_1, \cdots, j_p\}$ の場合，どのように順番を入れ

替えても両者は一致することはない．したがって，$\{1, \cdots, N\}$ の中から重複も許して $p$ 個とった $\mathbf{u}_{i_1\cdots i_p}$ は $S^p(X)$ の基底になっていることがわかる．ゆえに (1) が証明された．

(2)　$\Lambda^p(X)$ は $\{\mathcal{A}_p(\mathbf{u}_{i_1\cdots i_p}) : i_1, \cdots, i_p \in \{1, \cdots, N\}\}$ で生成されている．いま $i_1, \cdots, i_p$ の中に重複があった場合，適当な互換 $\tau$ により $P_\tau(\mathbf{u}_{i_1\cdots i_p}) = \mathbf{u}_{i_1\cdots i_p}$ となっている．したがって

$$\mathcal{A}_p(\mathbf{u}_{i_1\cdots i_p}) = \mathcal{A}_p P_\tau(\mathbf{u}_{i_1\cdots i_p}) = -\mathcal{A}_p(\mathbf{u}_{i_1\cdots i_p})$$

より $\mathcal{A}_p(\mathbf{u}_{i_1\cdots i_p}) = \mathbf{0}$ である．これより $p > N$ の場合 $\Lambda^p(X) = \{\mathbf{0}\}$ が得られる．$p \leq N$ の場合，$\{1, \cdots, N\}$ の中から相異なる $p$ 個をとった $\mathbf{u}_{i_1\cdots i_p}$ は $\Lambda^p(X)$ の基底になっていることがわかる．ゆえに (2) が得られる．■

**系 22. 17**　$1 \leq p \leq N$ の場合，

$$\{\mathcal{A}_p(\mathbf{u}_{i_1\cdots i_p}) : 1 \leq i_1 < \cdots < i_p \leq N\}$$

は $\Lambda^p(X)$ の基底になっている．

**問題 22. 8**　定理 22.15 の記号のもとに次のことを示せ．$\mathbf{t}$ が対称テンソルであるための必要十分条件は，任意の互換 $\sigma$ に対して，

$$t^{\sigma(i_1)\cdots\sigma(i_p)} = t^{i_1\cdots i_p}$$

となることである．また $\mathbf{t}$ が交代テンソルであるための必要十分条件は，任意の互換 $\sigma$ に対して，

$$t^{\sigma(i_1)\cdots\sigma(i_p)} = -\,t^{i_1\cdots i_p}$$

となることである．

## 22. 6　テンソル代数と外積代数

本節ではテンソル代数について述べる．まず多元環の定義からはじめる．

**定義 22. 5**　$A$ を $\boldsymbol{K}$ 上の線形空間とする．$m \in L(A, A; A)$ が $\mathbf{x}, \mathbf{y}, \mathbf{z} \in A$ に対して

$$m(m(\mathbf{x}, \mathbf{y}), \mathbf{z}) = m(\mathbf{x}, m(\mathbf{y}, \mathbf{z}))$$

をみたすとき，$(A, m)$ を $m$ を積とする $\boldsymbol{K}$ **上の多元環**という．混乱の生じない場

合は単に $A$ は $\boldsymbol{K}$ 上の**多元環である**という．積 $m(\mathbf{x}, \mathbf{y})$ はしばしば $\mathbf{x}\,\mathbf{y}$ と略記することも多い．ある $\mathbf{1} \in A$ で

$$m(\mathbf{1}, \mathbf{x}) = m(\mathbf{1}, \mathbf{x}) = \mathbf{x}$$

が成り立つものが存在するとき，$\mathbf{1}$ を $A$ の**単位元**という．このとき $A$ は単位元 $\mathbf{1}$ をもつという ( 注：第 2.3 節で導入した記号 $\mathbf{1}_N$ と混乱しないように ).

二つの多元環 $(A, m)$, $(A', m')$ に対して，$T \in L(A; A)$ かつ

$$T(m(\mathbf{x}, \mathbf{y})) = m'(T(\mathbf{x}), T(\mathbf{y})), \quad \mathbf{x}, \mathbf{y} \in A$$

をみたす $T$ を**準同形写像**という．特に準同形写像で全単射になっているものを**多元環の意味での同形写像**という．$(A, m)$, $(A', m')$ の間に多元環の意味での同形写像が存在するとき，この二つは**多元環の意味で同形**という．なおこれまで線形空間の意味での同形写像や同形，テンソル積の意味での同形写像や同形なども考えてきたが，前後の文脈からそれらと混乱を招くおそれのない場合は多元環の意味での同形写像，多元環の意味で同形をそれぞれ単に同形写像，同形と呼ぶ.

$A$ の部分空間 $B$ で，

$$\mathbf{x}, \mathbf{y} \in B \Longrightarrow m(\mathbf{x}, \mathbf{y}) \in B$$

をみたすとき $B$ を $A$ の**部分多元環**という.

$A_k\ (k = 0, 1, 2, \cdots)$ を $\boldsymbol{K}$ 上の線形空間とし，$A = \oplus_{k=0}^{\infty} A_k$ とする．$A$ が積 $m$ により多元環になっていて，さらに

$$\mathbf{x} \in A_j,\ \mathbf{y} \in A_k \Longrightarrow m(\mathbf{x}, \mathbf{y}) \in A_{j+k}$$

をみたすとき，$A$ を**次数つき多元環**という．また次数つき多元環 $A = \oplus_{k=0}^{\infty} A_k$ の部分環 $B$ が

$$B = \oplus_{k=0}^{\infty} (A_k \cap B)$$

をみたすとき**次数つき部分環**という.

### テンソル代数

テンソル積を用いて，テンソル代数と呼ばれる次数つき多元環を定義する．$X$ を $\boldsymbol{K}$ 上の $N$ 次元線形空間とする．このとき $T^k(X)$, $k = 0, 1, 2, \cdots$ の無限直和 $\oplus_{k=0}^{\infty} T^k(X)$ を考える (例 21.6 参照).

$$T(X) = \oplus_{k=0}^{\infty} T^k(X)$$

と表す．

$T^p(X)$ と $T^q(X)$ のテンソル積 $T^p(X)\otimes T^q(X)$ が定義できるが，このとき，$T^p(X)\otimes T^q(X)$ は $p+q$ 個の $X$ のテンソル積になっていて，

$$T^p(X)\otimes T^q(X)\approx T^{p+q}(X)$$

である．$\mathbf{x}\in T^p(X), \mathbf{y}\in T^q(X)$ のとき，$\mathbf{x}\otimes\mathbf{y}\in T^{p+q}(X)$ である．一般に $\mathbf{v}=\{\mathbf{v}_k\}_{k=0}^{\infty}$, $\mathbf{v}'=\{\mathbf{v}'_k\}_{k=0}^{\infty}\in\oplus_{k=0}^{\infty}T^k(X)$ に対して，

$$\mathbf{v}\otimes\mathbf{v}'=\sum_{p=0}^{\infty}\left(\sum_{r+s=p}\mathbf{v}_r\otimes\mathbf{v}'_s\right)\ \left(=\left\{\sum_{r+s=p}\mathbf{v}_r\otimes\mathbf{v}'_s\right\}_{p=0}^{\infty}\right)$$

と定義すると，$\mathbf{v}\otimes\mathbf{v}'\in T(X)$ である．特に $T^0(X)=\boldsymbol{K}\ni\mathbf{1}$ に対して，

$$\mathbf{1}\otimes\mathbf{v}=\mathbf{v}\otimes\mathbf{1}=\mathbf{v}$$

である．$(T(X),\otimes)$ は次数つき多元環になっている．この多元環 $T(X)$ を**テンソル代数**という．

**外積代数**

次に外積代数と呼ばれる次数つき多元環を定義する．$X$ を $\boldsymbol{K}$ 上の $N$ 次元線形空間とし，

$$\Lambda(X)=\left\{\sum_{k=0}^{\infty}\mathbf{v}_k\in T(X):\mathbf{v}_k\in\Lambda^k(X)\right\}$$

と定義する．$\Lambda^k(X)=\{\mathbf{0}\}, k>N$ であるから，$\sum_{k=0}^{\infty}\mathbf{v}_k\in\Lambda(X)$ ならば $\mathbf{v}_k=\mathbf{0}, k>N$ となっている．$\mathbf{v}=\sum_{k=0}^{\infty}\mathbf{v}_k$ $\mathbf{v}'=\sum_{k=0}^{\infty}\mathbf{v}'_k\in\Lambda(X)$ に対して，

$$\mathbf{v}\wedge\mathbf{v}'=\sum_{s=0}^{\infty}\left(\sum_{p+q=s}\frac{s!}{p!\,q!}\mathcal{A}_s\left(\mathbf{v}_p\otimes\mathbf{v}'_q\right)\right)$$

とする．これを $\mathbf{v}$ と $\mathbf{v}'$ の**外積**という (これまで $\wedge$ は二つの実数のうち大きくないほうを表す記号として用いてきた．今後は外積を表す記号としても用いるが，混乱は生じないであろう)．

$\mathbf{v}\wedge\mathbf{v}'\in\Lambda(X)$ である．特に $\mathbf{v}\in\Lambda^p(X), \mathbf{v}'\in\Lambda^q(X)$ の場合，

$$\mathbf{v}\wedge\mathbf{v}'=\frac{(p+q)!}{p!\,q!}\mathcal{A}_{p+q}\left(\mathbf{v}\otimes\mathbf{v}'\right).$$

次のことが成り立つ．

**定理 22.18**　$\mathbf{x}, \mathbf{y}, \mathbf{z} \in \Lambda(X)$ に対して,

$$(\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} = \mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}).$$

**証明.** $\mathbf{u} = \mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_p$, $\mathbf{v} = \mathbf{x}_{p+1} \otimes \cdots \otimes \mathbf{x}_{p+q}$ に対して

$$\begin{aligned} \mathcal{A}_{p+q}(\mathbf{u} \otimes \mathbf{v}) &= \mathcal{A}_{p+q}(\mathcal{A}_p(\mathbf{u}) \otimes \mathbf{v}) = \mathcal{A}_{p+q}(\mathbf{u} \otimes \mathcal{A}_q(\mathbf{v})) \\ &= \mathcal{A}_{p+q}(\mathcal{A}_p(\mathbf{u}) \otimes \mathcal{A}_q(\mathbf{v})) \end{aligned} \tag{22.8}$$

を示す．もしこれが示せれば,

$$\mathbf{x} = \mathcal{A}_p(\mathbf{u}),\ \mathbf{y} = \mathcal{A}_q(\mathbf{v}),\ \mathbf{w} = \mathbf{x}_{p+q+1} \otimes \cdots \otimes \mathbf{x}_{p+q+r},\ \mathbf{z} = \mathcal{A}_r(\mathbf{w})$$

に対して,

$$\begin{aligned} (\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} &= \frac{(p+q+r)!}{(p+q)!\,r'}\mathcal{A}_{p+q+r}\left(\frac{(p+q)!}{p!\,q!}\mathcal{A}_{p+q}(\mathbf{u} \otimes \mathbf{v}) \otimes \mathcal{A}_r(\mathbf{w})\right) \\ &= \frac{(p+q+r)!}{p!\,q!\,r!}\mathcal{A}_{p+q+r}(\mathcal{A}_{p+q}(\mathbf{u} \otimes \mathbf{v}) \otimes \mathcal{A}_r(\mathbf{w})) \\ &= \frac{(p+q+r)!}{p!\,q!\,r!}\mathcal{A}_{p+q+r}(\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}) \end{aligned}$$

が成り立つ．同様にして,

$$\mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}) = \frac{(p+q+r)!}{p!\,q!\,r!}\mathcal{A}_{p+q+r}(\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w})$$

も得られる．したがって定理がこの $\mathbf{x}, \mathbf{y}, \mathbf{z}$ の場合に証明された．一般の場合はテンソル積，$\mathcal{A}_s$ の $n$ 重線形性を用いて示される．以下では，(22.8) を証明する．まず

$$\begin{aligned} &\mathcal{A}_{p+q}(\mathcal{A}_p(\mathbf{u}) \otimes \mathbf{v}) = \mathcal{A}_{p+q}\left(\frac{1}{p!}\sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) P_\sigma(\mathbf{u}) \otimes \mathbf{v}\right) \\ &\quad = \mathcal{A}_{p+q}\left(\frac{1}{p!}\sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)(\mathbf{x}_{\sigma(1)} \otimes \cdots \otimes \mathbf{x}_{\sigma(p)}) \otimes (\mathbf{x}_{p+1} \otimes \cdots \otimes \mathbf{x}_{p+q})\right) \\ &\quad = \frac{1}{p!\,(p+q)!}\sum_{\tau \in S_{p+q}} \operatorname{sgn}(\tau) \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)\, \mathbf{x}_{\tau\sigma(1)} \otimes \\ &\qquad\qquad \cdots \otimes \mathbf{x}_{\tau\sigma(p)} \otimes \mathbf{x}_{\tau(p+1)} \otimes \cdots \otimes \mathbf{x}_{\tau(p+q)} \end{aligned} \tag{22.9}$$

が得られる．ここで $\tau$ が $S_{p+q}$ 全体にわたって動けば

$$\tau_\sigma = \begin{pmatrix} 1 & \cdots & p & p+1 & \cdots & p+q \\ \tau\sigma(1) & \cdots & \tau\sigma(p) & \tau(p+1) & \cdots & \tau(p+q) \end{pmatrix}$$

の形の置換も $S_{p+q}$ 全体を動く．さらに $\operatorname{sgn}(\tau_\sigma) = \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)$ であるから

(22.9)

$$
\begin{aligned}
&= \frac{1}{p!\,(p+q)!} \sum_{\tau \in S_{p+q}} p!\,\mathrm{sgn}\,(\tau)\, x_{\tau(1)} \otimes \cdots \otimes x_{\tau(p)} \otimes x_{\tau(p+1)} \otimes \cdots \otimes x_{\tau(p+q)} \\
&= \mathcal{A}_{p+q}\,(\mathbf{u} \otimes \mathbf{v})
\end{aligned}
$$

が得られる．同様にして

$$\mathcal{A}_{p+q}\,(\mathbf{u} \otimes \mathcal{A}_q\,(\mathbf{v})) = \mathcal{A}_{p+q}\,(\mathbf{u} \otimes \mathbf{v})$$

も示せる．$\mathbf{y} = \mathcal{A}_q(\mathbf{v})$ とおくと

$$
\begin{aligned}
\mathcal{A}_{p+q}\,(\mathcal{A}_p(\mathbf{u}) \otimes \mathcal{A}_q(\mathbf{v})) &= \mathcal{A}_{p+q}\,(\mathcal{A}_p(\mathbf{u}) \otimes \mathbf{y}) = \mathcal{A}_{p+q}\,(\mathbf{u} \otimes \mathbf{y}) \\
&= \mathcal{A}_{p+q}\,(\mathbf{u} \otimes \mathcal{A}_q(\mathbf{v}))
\end{aligned}
$$

が得られる．■

以上より，$\Lambda(X)$ は外積 $\wedge$ を積とした単位元 $\mathbf{1} \in \Lambda^0(X)$ をもつ次数つき多元環になっている．

**定義 22. 6**　$\mathbf{x}, \mathbf{y}, \mathbf{z} \in \Lambda(X)$ に対して，

$$\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} = (\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} = \mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z})$$

と定義する．帰納的に，$\mathbf{x}_1, \cdots, \mathbf{x}_n \in \Lambda(X)$ に対して，

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n = (\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_{n-1}) \wedge \mathbf{x}_n$$

と定義する．

本節の最後に外積のいくつかの基本的な性質を述べておく．

**定理 22. 19**　$\mathbf{x} \in T^p(X), \mathbf{y} \in T^q(X)$ のとき

$$\mathcal{A}_p(\mathbf{x}) \wedge \mathcal{A}_q(\mathbf{y}) = \frac{(p+q)!}{p!\,q!} \mathcal{A}_{p+q}\,(\mathbf{x} \otimes \mathbf{y})\,.$$

**証明．** (22.8) より

$$
\begin{aligned}
\mathcal{A}_p(\mathbf{x}) \wedge \mathcal{A}_q(\mathbf{y}) &= \frac{(p+q)!}{p!\,q!} \mathcal{A}_{p+q}\,(\mathcal{A}_p\,(\mathbf{x}) \otimes \mathcal{A}_q\,(\mathbf{y})) \\
&= \frac{(p+q)!}{p!\,q!} \mathcal{A}_{p+q}\,(\mathbf{x} \otimes \mathbf{y})\,.
\end{aligned}
$$

■

**系 22.20**　$\mathbf{x}_1, \cdots, \mathbf{x}_p \in X$ のとき

$$\mathcal{A}_p\left(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_p\right) = \frac{1}{p!}\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p.$$

**証明.**

$$\begin{aligned}
\mathcal{A}_p\left(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_p\right) &= \frac{(p-1)!}{p!}\mathcal{A}_{p-1}\left(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_{p-1}\right) \wedge \mathcal{A}_1\left(\mathbf{x}_p\right) \\
&= \frac{(p-1)!}{p!}\frac{(p-2)!}{(p-1)!}\mathcal{A}_{p-2}\left(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_{p-2}\right) \wedge \mathcal{A}_1\left(\mathbf{x}_{p-1}\right) \wedge \mathcal{A}_1\left(\mathbf{x}_p\right) \\
&= \frac{1}{p!}\mathcal{A}_1\left(\mathbf{x}_1\right) \wedge \cdots \wedge \mathcal{A}_1\left(\mathbf{x}_p\right) = \frac{1}{p!}\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p.
\end{aligned}$$

■

**定理 22.21**　$\mathbf{x} \in \Lambda^p(X)$, $\mathbf{y} \in \Lambda^q(X)$ のとき,

$$\mathbf{x} \wedge \mathbf{y} = (-1)^{pq}\ \mathbf{y} \wedge \mathbf{x}.$$

**証明.** $\tau \in S_{p+q}$ を $P_\tau\left(\mathbf{x} \otimes \mathbf{y}\right) = \mathbf{y} \otimes \mathbf{x}$ なるものとすると，$\mathrm{sgn}(\tau) = (-1)^{pq}$ であるから,

$$\begin{aligned}
\mathbf{x} \wedge \mathbf{y} = \mathcal{A}_p(\mathbf{x}) \wedge \mathcal{A}_q(\mathbf{y}) &= \frac{(p+q)!}{p!\,q!}\mathcal{A}_{p+q}\left(\mathbf{x} \otimes \mathbf{y}\right) \\
&= \frac{(p+q)!}{p!\,q!}\frac{1}{(p+q)!}\sum_{\sigma \in S_{p+q}} \mathrm{sgn}(\sigma)P_\sigma(\mathbf{x} \otimes \mathbf{y}) \\
&= \frac{(p+q)!}{p!\,q!}\frac{1}{(p+q)!}\sum_{\sigma \in S_{p+q}} \mathrm{sgn}(\tau\sigma)P_{\tau\sigma}(\mathbf{x} \otimes \mathbf{y}) \\
&= (-1)^{pq}\frac{(p+q)!}{p!\,q!}\frac{1}{(p+q)!}\sum_{\sigma \in S_{p+q}} \mathrm{sgn}(\sigma)P_\sigma(\mathbf{y} \otimes \mathbf{x}) \\
&= (-1)^{pq}\mathbf{y} \wedge \mathbf{x}.
\end{aligned}$$

■

**系 22.22**　$\mathbf{x}_1, \cdots, \mathbf{x}_p \in \Lambda^1(X) = X$ のとき，もしも $\mathbf{x}_i = \mathbf{x}_j$ となる $i, j$ が存在するならば

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_p = \mathbf{0}$$

である.

**証明.** $\mathbf{x}_i \wedge \mathbf{x}_j = -\mathbf{x}_j \wedge \mathbf{x}_i = -\mathbf{x}_i \wedge \mathbf{x}_j$ より $\mathbf{x}_i \wedge \mathbf{x}_j = \mathbf{0}$ である．これより系は

明らかに成り立つ．■

**問題 22.9** $X$ を $\boldsymbol{K}$ 上の有限次元線形空間とし，$A$ を $\mathbf{e}$ を単位元とする $\boldsymbol{K}$ 上の多元環とする．$\iota$ を $T^1(X)=X$ から $T(X)$ への自然な埋め込みとする (例 21.8 参照)．このとき任意の $S \in L(X;A)$ に対して $T(X)$ から $A$ への準同形写像 $\widetilde{T}$ で

$$\widetilde{T} \circ \iota = S, \quad \widetilde{T}\,(1) = \mathbf{e} \tag{22.10}$$

をみたすものがただ一つ存在する．

# 第23章　$k$-ベクトルと $k$-形式

本章を通して $X$ により $\boldsymbol{R}$ 上の $N$ 次元線形空間を表す．$\Lambda^k(X)$ の要素を $k$-ベクトル，$\Lambda^k(X^*)$ の要素を $k$-**形式**という．本章ではこれらについての基本事項を述べる．幾何学では $X$ は多様体の接空間，また $X^*$ は余接空間として使われる．

**ポイント 23**　$k$-ベクトル，$k$-形式は幾何学，数理物理学で重要な微分形式の理論の基礎をなすものである．また幾何学的測度論などでも用いられている．$k$-ベクトル，$k$-形式に馴染んでおくことは，これらの分野を学ぶ場合必要になることである．

## 23.1　$k$-ベクトルと線形空間の向き

本節では $k$-ベクトルの基本的な性質ならびにそれを用いて線形空間の「向き」という概念について述べる．

はじめに若干の記号を準備しておこう．$1 \leq k \leq N$ に対して

$$\mathcal{I}(k, N) = \left\{ I = (i_1, \cdots, i_k) \ : \ \begin{array}{l} 1 \leq i_1 < \cdots < i_k \leq N \\ i_1, \cdots, i_k \text{ は自然数} \end{array} \right\}$$

とおく．$I = (i_1, \cdots, i_k)$, $J = (j_1, \cdots, j_k) \in \mathcal{I}(k, N)$ に対して，$k$ 次正方行列

$$\begin{pmatrix} a^{i_1}_{j_1} & \cdots & a^{i_1}_{j_k} \\ \vdots & \ddots & \vdots \\ a^{i_k}_{j_1} & \cdots & a^{i_k}_{j_k} \end{pmatrix}$$

を $(a^I_J)$ と略記する．なお，ここで $a^{i_1}_{j_1}$ などは $i_1, j_1$ を添え字とするある実数を表しているもので，$a_{j_1}$ の $i_1$ 乗を表しているわけではない．

$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とする．$I = (i_1, \cdots, i_k) \in \mathcal{I}(k, N)$ に対して

$$\mathbf{u}_I = \mathbf{u}_{i_1} \wedge \cdots \wedge \mathbf{u}_{i_k}$$

とおく (これまで用いた記号 $\mathbf{u}_{\alpha_1 \cdots \alpha_k} = \mathbf{u}_{\alpha_1} \otimes \cdots \otimes \mathbf{u}_{\alpha_k}$ と混乱のないように注意)．このとき系 22.17 と系 22.20 より，$\{\mathbf{u}_I : I \in \mathcal{I}(k, N)\}$ が $\Lambda^k(X)$ の基底になっていることがわかる．したがって，任意の $\mathbf{x} \in \Lambda^k(X)$ に対して，

$$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \mathbf{u}_I$$

をみたす $\xi^I \in \boldsymbol{R}$, $I \in \mathcal{I}(k, N)$ が一意的に存在する (なお，ここで $\xi^I$ は $I$ を添え字とするある実数を表していて，$\xi^I = \xi^{i_1} \cdots \xi^{i_k}$ のような積を表しているわけではない)．

**定理 23.1** $\mathbf{x}_1, \cdots, \mathbf{x}_k \in X$ が $\mathbf{x}_j = \sum_{i=1}^{N} \xi_j^i \mathbf{u}_i$, $j = 1, \cdots, k$ と表されているとき，

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k = \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \xi_1^{i_1} & \cdots & \xi_k^{i_1} \\ \vdots & \ddots & \vdots \\ \xi_1^{i_k} & \cdots & \xi_k^{i_k} \end{pmatrix} \mathbf{u}_I.$$

**証明.**

$$\begin{aligned} \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k &= \sum_{i_1, \cdots, i_k = 1}^{N} \xi_1^{i_1} \cdots \xi_k^{i_k} \mathbf{u}_{i_1} \wedge \cdots \wedge \mathbf{u}_{i_k} \\ &= \sum_{I \in \mathcal{I}(k,N)} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \xi_1^{i_{\sigma(1)}} \cdots \xi_k^{i_{\sigma(k)}} \mathbf{u}_I \\ &= \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \xi_1^{i_1} & \cdots & \xi_k^{i_1} \\ \vdots & \ddots & \vdots \\ \xi_1^{i_k} & \cdots & \xi_k^{i_k} \end{pmatrix} \mathbf{u}_I. \end{aligned}$$

■

**系 23.2** $\mathbf{x}_1, \cdots, \mathbf{x}_k \in X$ とする．このとき，$\mathbf{x}_1, \cdots, \mathbf{x}_k$ が線形独立であるための必要十分条件は

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k \neq \mathbf{0}$$

なることである．

**証明.** 定理 23.1 での記号を用いる．$\mathbf{q}_i = \left(\xi_i^j\right)_{j=1}^{N} \in \boldsymbol{R}^N$, $i = 1, \cdots, k$ とおく．

定理 23.1 および定理 9.9 より，$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k \neq \mathbf{0}$ であるための必要十分条件は $\mathbf{q}_1, \cdots, \mathbf{q}_k$ が線形独立なことである．

$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k \neq \mathbf{0}$ とする．$\alpha^1 \mathbf{x}_1 + \cdots + \alpha^k \mathbf{x}_k = \mathbf{0}$ とすると，

$$\alpha^1 \mathbf{x}_1 + \cdots + \alpha^k \mathbf{x}_k = \sum_{j=1}^{k} \alpha^j \left( \sum_{i=1}^{N} \xi_j^i \mathbf{u}_i \right) = \sum_{i=1}^{N} \left( \sum_{j=1}^{k} \xi_j^i \alpha^j \right) \mathbf{u}_i$$

より $\sum_{j=1}^{k} \xi_j^i \alpha^j = 0,\ i = 1, \cdots, N$ が成り立つ．ゆえに

$$\alpha^1 \mathbf{q}_1 + \cdots + \alpha^k \mathbf{q}_k = \mathbf{0}$$

となるが，$\mathbf{q}_1, \cdots, \mathbf{q}_k$ が線形独立であることより $\alpha^1 = \cdots = \alpha^k = 0$ を得る．ゆえに $\mathbf{x}_1, \cdots, \mathbf{x}_k$ が線形独立である．逆に $\mathbf{x}_1, \cdots, \mathbf{x}_k$ が線形独立であるとする．$\alpha^1 \mathbf{q}_1 + \cdots + \alpha^k \mathbf{q}_k = \mathbf{0}$ であるとすると，前半の議論を逆にたどって $\alpha^1 \mathbf{x}_1 + \cdots + \alpha^k \mathbf{x}_k = \mathbf{0}$ が得られる．ゆえに $\alpha^1 = \cdots = \alpha^k = 0$ となり，$\mathbf{q}_1, \cdots, \mathbf{q}_k$ が線形独立であることが導かれる．よって $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k \neq \mathbf{0}$ である．■

さて，$k$-ベクトル $\mathbf{x}$ は二つの基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ および $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ により，二つの表示

$$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \mathbf{u}_J, \quad \mathbf{x} = \sum_{J \in \mathcal{I}(k,N)} \widetilde{\xi}^J \widetilde{\mathbf{u}}_J$$

をもつ．このときの係数 $\xi^I$ と $\widetilde{\xi}^J$ の間の関係式を導いておく．

**定理 23. 3**　(21.1) の記号を用いて $\left(u_j^i\right)^{-1} = \left(\widetilde{u}_j^i\right)$ とおく．次が成り立つ．

$$\widetilde{\xi}^I = \sum_{J \in \mathcal{I}(k,N)} \det\left(\widetilde{u}_J^I\right) \xi^J,$$

$$\xi^I = \sum_{J \in I(k,N)} \det\left(u_J^I\right) \widetilde{\xi}^J.$$

**証明.** $\mathbf{u}_j = \sum_{i=1}^{N} \widetilde{u}_j^i \widetilde{\mathbf{u}}_i$ であるから

$$\sum_{I \in \mathcal{I}(k,N)} \widetilde{\xi}^I \widetilde{\mathbf{u}}_I = \sum_{J \in \mathcal{I}(k,N)} \xi^J \mathbf{u}_J = \sum_{J \in \mathcal{I}(k,N)} \xi^J \sum_{I \in \mathcal{I}(k,N)} \det\left(\widetilde{u}_J^I\right) \widetilde{\mathbf{u}}_I$$

$$= \sum_{I \in \mathcal{I}(k,N)} \left( \sum_{J \in \mathcal{I}(k,N)} \det\left(\widetilde{u}_J^I\right) \xi^J \right) \widetilde{\mathbf{u}}_I.$$

■

**定理 23.4** $\mathbf{x}_1, \cdots, \mathbf{x}_k \in X$, $\mathbf{g}^1, \cdots, \mathbf{g}^k \in X^*$ に対して

$$(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k)\left(\mathbf{g}^1, \cdots, \mathbf{g}^k\right) = \det \begin{pmatrix} \mathbf{g}^1(\mathbf{x}_1) & \cdots & \mathbf{g}^1(\mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ \mathbf{g}^k(\mathbf{x}_1) & \cdots & \mathbf{g}^k(\mathbf{x}_k) \end{pmatrix}.$$

**証明.** 定義より

$$\begin{aligned} \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k &= p!\,\mathcal{A}_k(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k) \\ &= \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \mathbf{x}_{\sigma(1)} \otimes \cdots \otimes \mathbf{x}_{\sigma(k)} \end{aligned}$$

であるから

$$\begin{aligned} (\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k)\left(\mathbf{g}^1, \cdots, \mathbf{g}^k\right) &= \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \mathbf{x}_{\sigma(1)} \otimes \cdots \otimes \mathbf{x}_{\sigma(k)}\left(\mathbf{g}^1, \cdots, \mathbf{g}^k\right) \\ &= \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \mathbf{g}^1\left(\mathbf{x}_{\sigma(1)}\right) \cdots \mathbf{g}^k\left(\mathbf{x}_{\sigma(k)}\right) \\ &= \det \begin{pmatrix} \mathbf{g}^1(\mathbf{x}_1) & \cdots & \mathbf{g}^1(\mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ \mathbf{g}^k(\mathbf{x}_1) & \cdots & \mathbf{g}^k(\mathbf{x}_k) \end{pmatrix}. \end{aligned}$$

■

$\left\{\mathbf{f}^1, \cdots, \mathbf{f}^N\right\}$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ の双対基底とする．これより，次のことがわかる．

**系 23.5** $I \in \mathcal{I}(k, N)$ と $\beta_1, \cdots, \beta_k \in \{1, 2, \cdots, N\}$ に対して

$$\mathbf{u}_I\left(\mathbf{f}^{\beta_1}, \cdots, \mathbf{f}^{\beta_k}\right) = \begin{cases} 0, & \{i_1, \cdots, i_k\} \neq \{\beta_1, \cdots, \beta_k\} \\ \operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_k \\ \beta_1 & \cdots & \beta_k \end{pmatrix} & \{i_1, \cdots, i_k\} = \{\beta_1, \cdots, \beta_k\} \end{cases}.$$

**証明.**

$$\begin{aligned} \mathbf{u}_\alpha\left(\mathbf{f}^{\beta_1}, \cdots, \mathbf{f}^{\beta_k}\right) &= (\mathbf{u}_{i_1} \wedge \cdots \wedge \mathbf{u}_{i_k})\left(\mathbf{f}^{\beta_1}, \cdots, \mathbf{f}^{\beta_k}\right) \\ &= \det \begin{pmatrix} \delta_{i_1}^{\beta_1} & \cdots & \delta_{i_k}^{\beta_1} \\ \vdots & \ddots & \vdots \\ \delta_{i_1}^{\beta_k} & \cdots & \delta_{i_k}^{\beta_k} \end{pmatrix} \end{aligned}$$

$$= \begin{cases} 0, & \{i_1, \cdots, i_k\} \neq \{\beta_1, \cdots, \beta_k\} \\ \operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_k \\ \beta_1 & \cdots & \beta_k \end{pmatrix} & \{i_1, \cdots, i_k\} = \{\beta_1, \cdots, \beta_k\} \end{cases}.$$

■

本節の最後に線形空間 $X$ に「向き」を定義する．

まず $X$ の基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を一つ選び固定する．この操作を $X$ の**向きを定める**という．あるいは基底 $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ により $X$ は**正に向き付けられている**という．

いま $X$ の別の基底 $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ をとってくる．系 23.2 より

$$\omega = \mathbf{u}_1 \wedge \cdots \wedge \mathbf{u}_N,$$
$$\widetilde{\omega} = \widetilde{\mathbf{u}}_1 \wedge \cdots \wedge \widetilde{\mathbf{u}}_N$$

はともに $\mathbf{0}$ ではない $N$-ベクトルである．また定理 22.16 より $\dim \Lambda^N(X) = 1$ であるから $\omega = t\widetilde{\omega}$ なる実数 $t$ が存在する．$\omega$ も $\widetilde{\omega}$ も $\mathbf{0}$ ではないので，$t > 0$ か $t < 0$ である．$t > 0$ の場合，基底 $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ は**正の向きをもつ基底**という．また $t < 0$ の場合は**負の向きをもつ基底**という．

$\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ を $X$ の基底とし，$\widetilde{\mathbf{u}}_j = \sum_{i=1}^{N} u_j^i \mathbf{u}_i$ が成り立っているとする．$I \in \mathcal{I}(k, N)$ に対して，$\widetilde{\mathbf{u}}_I = \widetilde{\mathbf{u}}_{i_1} \wedge \cdots \wedge \widetilde{\mathbf{u}}_{i_k}$ とおく．定理 23.1 より次が成り立つ．

**系 23.6**　$I \in \mathcal{I}(k, N)$ に対して，

$$\begin{aligned} \widetilde{\mathbf{u}}_I &= \sum_{J \in \mathcal{I}(k,N)} \det \begin{pmatrix} u_{i_1}^{j_1} & \cdots & u_{i_k}^{j_1} \\ \vdots & \ddots & \vdots \\ u_{i_1}^{j_k} & \cdots & u_{i_k}^{j_k} \end{pmatrix} \mathbf{u}_{j_1} \wedge \cdots \wedge \mathbf{u}_{j_k} \\ &= \sum_{J \in \mathcal{I}(k,N)} \det\left(u_I^J\right) \mathbf{u}_J. \end{aligned}$$

この系より，特に

$$\widetilde{\omega} = \det \begin{pmatrix} u_1^1 & \cdots & u_N^1 \\ \vdots & \ddots & \vdots \\ u_1^N & \cdots & u_N^N \end{pmatrix} \omega$$

であるから，もし $\det\left(u_j^i\right) > 0$ ならば $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ は正に向き付けられていて，$\det\left(u_j^i\right) < 0$ ならば $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ は負に向き付けられている．

## 23.2 $k$-形式

本節では $k$-形式の基礎事項を述べる．以下では $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とし，$\left\{\mathbf{f}^1, \cdots, \mathbf{f}^N\right\}$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ の双対基底とする．$I \in \mathcal{I}(k, N)$ に対して

$$\mathbf{f}^I = \mathbf{f}^{i_1} \wedge \cdots \wedge \mathbf{f}^{i_k}$$

とすると，系 22.17 と系 22.20 より，$\left\{\mathbf{f}^I : I \in \mathcal{I}(k, N)\right\}$ が $\Lambda^k(X^*)$ の基底になっていることがわかる．$k$-形式 $\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k$ は基底 $\mathbf{f}^I$ を用いて次のように表せる．

**定理 23.7**　$\mathbf{g}^j = \sum_{i=1}^{N} \eta_i^j \mathbf{f}^i$, $j = 1, \cdots, k$ となっているとき，

$$\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k = \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \eta_{i_1}^1 & \cdots & \eta_{i_k}^1 \\ \vdots & \ddots & \vdots \\ \eta_{i_1}^k & \cdots & \eta_{i_k}^k \end{pmatrix} \mathbf{f}^I.$$

**証明.**

$$\begin{aligned}
\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k &= \sum_{i_1, \cdots, i_k = 1}^{N} \eta_{i_1}^1 \cdots \eta_{i_k}^k \mathbf{f}^{i_1} \wedge \cdots \wedge \mathbf{f}^{i_k} \\
&= \sum_{1 \le i_1 < \cdots < i_k \le N} \sum_{\sigma \in S_k} \eta_{i_{\sigma(1)}}^1 \cdots \eta_{i_{\sigma(k)}}^k \mathbf{f}^{i_{\sigma(1)}} \wedge \cdots \wedge \mathbf{f}^{i_{\sigma(k)}} \\
&= \sum_{1 \le i_1 < \cdots < i_k \le N} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \eta_{i_{\sigma(1)}}^1 \cdots \eta_{i_{\sigma(k)}}^k \mathbf{f}^{i_1} \wedge \cdots \wedge \mathbf{f}^{i_k} \\
&= \sum_{1 \le i_1 < \cdots < i_k \le N} \det \begin{pmatrix} \eta_{i_1}^1 & \cdots & \eta_{i_k}^1 \\ \vdots & \ddots & \vdots \\ \eta_{i_1}^k & \cdots & \eta_{i_k}^k \end{pmatrix} \mathbf{f}^{i_1} \wedge \cdots \wedge \mathbf{f}^{i_k}.
\end{aligned}$$

■

**注意 23.1**　$\mathbf{g}^1, \cdots, \mathbf{g}^k \in X^*$ と $\tau_1, \cdots, \tau_k \in (X^*)^*$ に対して，

$$\left(\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k\right)(\tau_1, \cdots, \tau_k) = \det \begin{pmatrix} \tau_1(\mathbf{g}^1) & \cdots & \tau_k(\mathbf{g}^1) \\ \vdots & \ddots & \vdots \\ \tau_1(\mathbf{g}^k) & \cdots & \tau_k(\mathbf{g}^k) \end{pmatrix}$$

であるが，ここである $\mathbf{x}_j \in X$ で $\tau_j = \tau_{\mathbf{x}_j}$ となるものが一意的に存在する．したがって，

$$\left(\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k\right)(\tau_1, \cdots, \tau_k) = \det \begin{pmatrix} \mathbf{g}^1(\mathbf{x}_1) & \cdots & \mathbf{g}^1(\mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ \mathbf{g}^k(\mathbf{x}_1) & \cdots & \mathbf{g}^k(\mathbf{x}_k) \end{pmatrix}$$
$$= \left(\mathbf{g}^1 \wedge \cdots \wedge \mathbf{g}^k\right)(\mathbf{x}_1, \cdots, \mathbf{x}_k)$$

と考えることができる．

$\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ を $X$ の基底とし，$\left\{\widetilde{\mathbf{f}}^1, \cdots, \widetilde{\mathbf{f}}^N\right\}$ を $\{\widetilde{\mathbf{u}}_1, \cdots, \widetilde{\mathbf{u}}_N\}$ の双対基底とし，(21.1), (21.7) の関係式が成り立っているものとする．このとき $k$-ベクトルの場合と同様にして

$$\mathbf{f}^I = \sum_{J \in \mathcal{I}(k,N)} \det\left(u_J^I\right) \widetilde{\mathbf{f}}^J,$$
$$\widetilde{\mathbf{f}}^I = \sum_{J \in \mathcal{I}(k,N)} \det\left(\widetilde{u}_J^I\right) \mathbf{f}^J$$

と表せることがわかる．これより次の変換則が証明できる．

**定理 23.8**　$\mathbf{g} = \sum_{I \in \mathcal{I}(k,N)} \eta_I \mathbf{f}^I = \sum_{I \in \mathcal{I}(k,N)} \widetilde{\eta}_I \widetilde{\mathbf{f}}^I$ とすると，

$$\widetilde{\eta}_I = \sum_{J \in \mathcal{I}(k,N)} \det\left(u_I^J\right) \eta_J,$$
$$\eta_I = \sum_{I \in \mathcal{I}(k,N)} \det\left(\widetilde{u}_I^J\right) \widetilde{\eta}_J.$$

**証明.**

$$\sum_{I \in \mathcal{I}(k,N)} \widetilde{\eta}_I \widetilde{\mathbf{f}}^I = \sum_{J \in \mathcal{I}(k,N)} \eta_J \mathbf{f}^J = \sum_{J \in \mathcal{I}(k,N)} \eta_J \sum_{I \in \mathcal{I}(k,N)} \det\left(u_I^J\right) \widetilde{\mathbf{f}}^I$$
$$= \sum_{I \in \mathcal{I}(k,N)} \left( \sum_{J \in \mathcal{I}(k,N)} \det\left(u_I^J\right) \eta_J \right) \widetilde{\mathbf{f}}^I.$$

■

**定義 23.1**　$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \mathbf{u}_I \in \Lambda^k(X)$, $\mathbf{h} = \sum_{I \in \mathcal{I}(k,N)} \eta_I \mathbf{f}^I \in \Lambda^k(X^*)$ に対して，

$$\langle \mathbf{x}, \mathbf{h} \rangle_{*,k} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \eta_I \tag{23.1}$$

とする．混乱のない場合は $\langle \mathbf{x}, \mathbf{h} \rangle_{*,k}$ を $\langle \mathbf{x}, \mathbf{h} \rangle_*$ と略記する．

この定義が意味をもつためには，(23.1) の右辺の値が基底のとり方によらず一定であることを証明しておかねばならない．(23.1) の右辺の値が基底のとり方によらないことは次のように証明される．$\mathbf{x}, \mathbf{h}$ が別の基底を用いて

$$\mathbf{x} = \sum_{I\in\mathcal{I}(k,N)} \widetilde{\xi}^I \widetilde{\mathbf{u}}_I, \quad \mathbf{h} = \sum_{I\in\mathcal{I}(k,N)} \widetilde{\eta}_I \widetilde{\mathbf{f}}^I$$

と表されているとする．このとき定理 23.8 より

$$\begin{aligned}\sum_{I\in\mathcal{I}(k,N)} \xi^I \eta_I &= \sum_{I\in\mathcal{I}(k,N)} \sum_{J\in\mathcal{I}(k,N)} \det\left(u_J^I\right) \widetilde{\xi}^J \sum_{J'\in I(k,N)} \det\left(\widetilde{u}_I^{J'}\right) \widetilde{\eta}_{J'} \\ &= \sum_{J\in\mathcal{I}(k,N)} \sum_{J'\in\mathcal{I}(k,N)} \left(\sum_{I\in\mathcal{I}(k,N)} \det\left(u_J^I\right) \det\left(\widetilde{u}_I^{J'}\right)\right) \widetilde{\xi}^J \widetilde{\eta}_{J'}.\end{aligned}$$

ここで定理 7.6 を用いれば $J, J' \in \mathcal{I}(k,N)$ に対して

$$\begin{aligned}\sum_{I\in\mathcal{I}(k,N)} \det\left(\widetilde{u}_I^{J'}\right) \det\left(u_J^I\right) &= \det\left(\begin{pmatrix} \widetilde{u}_1^{j_1'} & \cdots & \widetilde{u}_N^{j_1'} \\ \vdots & \ddots & \vdots \\ \widetilde{u}_1^{j_k'} & \cdots & \widetilde{u}_N^{j_k'} \end{pmatrix} \begin{pmatrix} u_{j_1}^1 & \cdots & u_{j_k}^1 \\ \vdots & \ddots & \vdots \\ u_{j_1}^N & \cdots & u_{j_k}^N \end{pmatrix}\right) \\ &= \det\begin{pmatrix} \sum_{i=1}^N \widetilde{u}_i^{j_1'} u_{j_1}^i & \cdots & \sum_{i=1}^N \widetilde{u}_i^{j_1'} u_{j_k}^i \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^N \widetilde{u}_i^{j_k'} u_{j_1}^i & \cdots & \sum_{i=1}^N \widetilde{u}_i^{j_k'} u_{j_k}^i \end{pmatrix} \\ &= \det\begin{pmatrix} \delta_{j_1}^{j_1'} & \cdots & \delta_{j_k}^{j_1'} \\ \vdots & \ddots & \vdots \\ \delta_{j_1}^{j_k'} & \cdots & \delta_{j_k}^{j_k'} \end{pmatrix} \\ &= \begin{cases} 1, & j_1 = j_1', \cdots, j_k = j_k' \\ 0, & \text{その他} \end{cases} .\end{aligned}$$

ゆえに

$$\sum_{I\in\mathcal{I}(k,N)} \xi^I \eta_I = \sum_{I\in\mathcal{I}(k,N)} \widetilde{\xi}^I \widetilde{\eta}_I$$

が得られる．

定義 23.1 で定義された $\langle\cdot,\cdot\rangle_{*,k}$ を用いて，$\Lambda^k(X^*)$ から $\left(\Lambda^k(X)\right)^*$ への線形空間の意味での同形写像を構成することができる．まず $\mathbf{h} \in \Lambda^k(X^*)$ を任意にとり

固定したとき，$\mathbf{x} \in \Lambda^k(X)$ に対して，$\langle \mathbf{x}, \mathbf{h} \rangle_{*,k} \in \boldsymbol{R}$ を対応させる写像を $\langle \cdot, \mathbf{h} \rangle_{*,k}$ と表すと，$\langle \cdot, \mathbf{h} \rangle_{*,k} \in \left(\Lambda^k(X)\right)^*$ であることに注意しておく．

**定理 23.9**　$1 \le k \le N$ を自然数とする．

$$\Gamma : \Lambda^k(X^*) \ni \mathbf{h} \longmapsto \langle \cdot, \mathbf{h} \rangle_{*,k} \in \left(\Lambda^k(X)\right)^*$$

とする．このとき $\Gamma$ は $\Lambda^k(X^*)$ から $\left(\Lambda^k(X)\right)^*$ への線形空間の意味での同形写像で，任意の $\varphi \in \left(\Lambda^k(X)\right)^*$ に対して，

$$\varphi(\mathbf{x}) = \left\langle \mathbf{x}, \Gamma^{-1}(\varphi) \right\rangle_{*,k}, \quad \mathbf{x} \in \Lambda^k(X)$$

をみたしている．

**証明.**　明らかに $\Gamma$ は線形写像である．全射であることを示す．任意に $\varphi \in \left(\Lambda^k(X)\right)^*$ をとる．$\mathbf{h}_\varphi = \sum_{I \in \mathcal{I}(k,N)} \varphi(\mathbf{u}_I)\,\mathbf{h}^I$ とすると，$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \mathbf{u}_I$ に対して

$$\langle \mathbf{x}, \mathbf{h}_\varphi \rangle_{*,k} = \sum_{I \in \mathcal{I}(k,N)} \xi^I \varphi(\mathbf{u}_I) = \varphi\left(\sum_{I \in \mathcal{I}(k,N)} \xi^I \mathbf{u}_I\right) = \varphi(\mathbf{x})$$

である．ゆえに $\Gamma(\mathbf{h}_\varphi) = \varphi$ である．

次に単射であることを示すが，これは任意の $\mathbf{x} \in \Lambda^p(X)$ に対して $\langle \mathbf{x}, \mathbf{h} \rangle_{*,k} = 0$ ならば $\mathbf{h} = \mathbf{0}$ であることを示すのと同値である．この主張は，たとえば $\mathbf{h} = \sum_{I \in \mathcal{I}(k,N)} \xi_I \mathbf{f}^I$ とするとき，$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} \xi_I \mathbf{u}_I \in \Lambda^p(X)$ に対して

$$0 = \langle \mathbf{x}, \mathbf{h} \rangle_{*,k} = \sum_{I \in \mathcal{I}(k,N)} |\xi_I|^2$$

となることによる．ゆえに $T$ が同形写像であることが示された．定理の後半の主張は，前半の主張より明らかである．■

**注意 23.2**　$\mathbf{h} \in \Lambda^k(X^*)$ に対して

$$\Gamma(\mathbf{h})(\mathbf{x}) = \langle \mathbf{x}, \mathbf{h} \rangle_{*,k}, \quad \mathbf{x} \in \Lambda^k(X)$$

であるが，特に $k = 1$ の場合，$\mathbf{h} \in X^*$ に対して $\Gamma(\mathbf{h}) = \langle \cdot, \mathbf{h} \rangle_{*,1} \in X^*$ であるから，

$$\Gamma(\mathbf{h})(\mathbf{x}) = \langle \mathbf{x}, \mathbf{h} \rangle_{*,1} = \mathbf{h}(\mathbf{x}), \quad \mathbf{x} \in X$$

となっている．したがって $k = 1$ の場合は $\Gamma = Id_{X^*}$ である．

**定理 23.10** $\mathbf{x}_j = \sum_{n=1}^{N} \xi_j^n \mathbf{u}_n \in X$, $\mathbf{h}^j = \sum_{n=1}^{N} \eta_n^j \mathbf{f}^n \in X^*$, $j = 1, \cdots, k$ とすると

$$\left\langle \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k, \mathbf{h}^1 \wedge \cdots \wedge \mathbf{h}^k \right\rangle_{*,p} = \det \begin{pmatrix} \langle \mathbf{x}_1, \mathbf{h}^1 \rangle_{*,1} & \cdots & \langle \mathbf{x}_k, \mathbf{h}^1 \rangle_{*,1} \\ \vdots & \ddots & \vdots \\ \langle \mathbf{x}_1, \mathbf{h}^k \rangle_{*,1} & \cdots & \langle \mathbf{x}_k, \mathbf{h}^k \rangle_{*,1} \end{pmatrix}$$

$$= \det \begin{pmatrix} \mathbf{h}^1 (\mathbf{x}_1) & \cdots & \mathbf{h}^1 (\mathbf{x}_k) \\ \vdots & \ddots & \vdots \\ \mathbf{h}^k (\mathbf{x}_1) & \cdots & \mathbf{h}^k (\mathbf{x}_k) \end{pmatrix}.$$

**証明.** 定理 23.1, 23.7 より

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k = \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \xi_1^{i_1} & \cdots & \xi_k^{i_1} \\ \vdots & \ddots & \vdots \\ \xi_1^{i_k} & \cdots & \xi_k^{i_k} \end{pmatrix} \mathbf{u}_I,$$

$$\mathbf{h}^1 \wedge \cdots \wedge \mathbf{h}^k = \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \eta_{i_1}^1 & \cdots & \eta_{i_k}^1 \\ \vdots & \ddots & \vdots \\ \eta_{i_k}^k & \cdots & \eta_{i_k}^k \end{pmatrix} \mathbf{f}^I$$

である．したがって，

$$\begin{aligned} &\left\langle \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k, \mathbf{h}^1 \wedge \cdots \wedge \mathbf{h}^k \right\rangle_* \\ &= \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} \xi_1^{i_1} & \cdots & \xi_k^{i_1} \\ \vdots & \ddots & \vdots \\ \xi_1^{i_k} & \cdots & \xi_k^{i_k} \end{pmatrix} \det \begin{pmatrix} \eta_{i_1}^1 & \cdots & \eta_{i_p}^1 \\ \vdots & \ddots & \vdots \\ \eta_{i_1}^k & \cdots & \eta_{i_p}^k \end{pmatrix} \\ &=: (I) \end{aligned}$$

であるが，定理 7.6 より，

$$(I) = \det \left( \begin{pmatrix} \eta_1^1 & \cdots & \eta_N^1 \\ \vdots & \ddots & \vdots \\ \eta_1^k & \cdots & \eta_N^k \end{pmatrix} \begin{pmatrix} \xi_1^1 & \cdots & \xi_k^1 \\ \vdots & \ddots & \vdots \\ \xi_1^N & \cdots & \xi_k^N \end{pmatrix} \right)$$

$$
= \det \begin{pmatrix} \sum_{n=1}^{N} \eta_n^1 \xi_1^n & \cdots & \sum_{n=1}^{N} \eta_n^1 \xi_k^n \\ \vdots & \ddots & \vdots \\ \sum_{n=1}^{N} \eta_n^k \xi_1^n & \cdots & \sum_{n=1}^{N} \eta_n^k \xi_k^n \end{pmatrix}
$$

となる．$\left\langle \mathbf{x}_j, \mathbf{h}^k \right\rangle_{*,1} = \sum_{n=1}^{N} \eta_n^k \xi_j^n$ であるから定理が証明された．■

## 23.3　$k$-ベクトルに対する内積

$X$ を $\boldsymbol{R}$ 上の $N$ 次元線形空間で，$X$ 上に内積 $\langle \cdot, \cdot \rangle$ が定義されているものとする．本節では，この $X$ 上に定義された内積から $\Lambda^k(X)$ 上の内積を定義する．

系 21.15 より，$\iota$ は $X$ から $X^*$ への同形になっていて，さらに注意 23.1 より，$\mathbf{x}, \mathbf{a} \in X$ に対して

$$
\langle \mathbf{x}, \iota_{\mathbf{a}} \rangle_{*,1} = \iota_{\mathbf{a}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{a} \rangle
$$

が成り立っていることを想起しておく．

$\eta \in \Lambda^k(X)$ に対して，$\iota_\sharp(\eta)$ を

$$
\iota_\sharp(\eta)(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k) = \langle \eta, \iota_{\mathbf{x}_1} \wedge \cdots \wedge \iota_{\mathbf{x}_k} \rangle_{*,k}, \quad \mathbf{x}_1, \cdots, \mathbf{x}_k \in X
$$

により定める．$\Lambda^k(X)$ は $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k$ の形の $k$-ベクトルの線形結合で表されているから，$\iota_\sharp(\eta)$ を $\Lambda^k(X)$ 上で定義された線形汎関数に拡張できる．このとき $\iota_\sharp$ により $\eta \in \Lambda^k(X)$ を $\iota_\sharp(\eta) \in \left(\Lambda^k(X)\right)^*$ に対応させる写像を表すことにすると，$\iota_\sharp \in L\left(\Lambda^k(X); \left(\Lambda^k(X)\right)^*\right)$ となる．$\Lambda^k(X) \ni \xi, \eta$ に対して，

$$
\langle \xi, \eta \rangle_k = \iota_\sharp(\eta)(\xi)
$$

と定める．後述の定理 23.11 より容易に $\langle \cdot, \cdot \rangle_k$ は $\Lambda^k(X)$ 上の内積になっていることがわかる．$\langle \cdot, \cdot \rangle_k$ を **$X$ の内積から誘導された $\Lambda^k(X)$ 上の内積**という．

**注意 23.3**　特に $k = 1$ の場合は

$$
\langle \xi, \eta \rangle_1 = \iota_\sharp(\eta)(\xi) = \langle \eta, \iota_\xi \rangle_{*,1} = \langle \eta, \xi \rangle .
$$

すなわち

$$
\langle \cdot, \cdot \rangle_1 = \langle \cdot, \cdot \rangle
$$

が成り立っている．

$\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の基底とする．このとき，$k$-ベクトルは

$$\mathbf{x} = \sum_{I \in \mathcal{I}(k,N)} x^I \mathbf{u}_I, \qquad \mathbf{y} = \sum_{I \in \mathcal{I}(k,N)} y^I \mathbf{u}_I$$

のように表せているが，このとき得られる係数 $x^I$, $y^I$ を用いて $\langle \mathbf{x}, \mathbf{y} \rangle_k$ を計算する方法を述べる．まず $i, j \in \{1, \cdots, N\}$ に対して

$$g_{ij} = \langle \mathbf{u}_i, \mathbf{u}_j \rangle$$

とおく．このとき内積の定義から $N$ 次正方行列 $(g_{ij})$ は正定値実対称行列になっている．したがってこの行列は正則であり，以下では $\left(g^{ij}\right)$ により $(g_{ij})$ の逆行列を表す．すなわち $\delta^i_j = \delta_{ij}$ と表せば

$$\sum_{k=1}^{N} g^{ik} g_{kj} = \delta^i_j$$

が成り立っている．

$I, J \in \mathcal{I}(k, N)$ に対して

$$g_{IJ} = g_{i_1 j_1} \cdots g_{i_k j_k},$$
$$g^{IJ} = g^{i_1 j_1} \cdots g^{i_k j_k}$$

とおく．また

$$\delta^I_J = \delta^{i_1}_{j_1} \cdots \delta^{i_k}_{j_k}$$

とおく．次のことが成り立つ．

**定理 23.11**　$k$-ベクトル $\mathbf{x} = \sum\limits_{I \in \mathcal{I}(k,N)} x^I \mathbf{u}_I$, $\mathbf{y} = \sum\limits_{I \in \mathcal{I}(k,N)} y^I \mathbf{u}_I$ に対して

$$\langle \mathbf{x}, \mathbf{y} \rangle_k = \sum_{I,J \in \mathcal{I}(k,N)} g_{IJ} x^I y^J.$$

この定理を証明するために次の補題を示しておく．

**補題 23.12**　$i = 1, \cdots, N$ に対して

$$\iota_{\mathbf{u}_i} = \sum_{j=1}^{N} g_{ij} \mathbf{f}^j.$$

**証明.** 任意の $\mathbf{x} = \sum\limits_{j=1}^{N} x^j \mathbf{u}_j \in X$ に対して

$$\iota_{\mathbf{u}_i}(\mathbf{x}) = \sum_{j=1}^{N} x^j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \sum_{j=1}^{N} g_{ij} x^j$$

であるが，一方

$$\begin{aligned}\sum_{j=1}^{N} g_{ij} \mathbf{f}^j(\mathbf{x}) &= \sum_{j=1}^{N} g_{ij} \mathbf{f}^j \left( \sum_{l=1}^{N} x^l \mathbf{u}_l \right) = \sum_{j=1}^{N} \sum_{l=1}^{N} g_{ij} \delta_l^j x^l \\ &= \sum_{j=1}^{N} g_{ij} x^j\end{aligned}$$

となっている．■

**補題 23.13** $I, J \in \mathcal{I}(k, N)$ に対して

$$\iota_\sharp(\mathbf{u}_I)(\mathbf{u}_J) = g_{IJ}.$$

**証明.** $\iota_\sharp$ の定義と補題 23.12 より

$$\begin{aligned}\iota_\sharp(\mathbf{u}_I)(\mathbf{u}_J) &= \left\langle \mathbf{u}_I, \iota_{\mathbf{u}_{j_1}} \wedge \cdots \wedge \iota_{\mathbf{u}_{j_k}} \right\rangle_{*,k} = \sum_{K \in \mathcal{I}(k,N)} g_{JK} \left\langle \mathbf{u}_I, f^K \right\rangle_{*,k} \\ &= \sum_{K \in \mathcal{I}(k,N)} g_{JK} \delta_I^K = g_{JI} = g_{IJ}.\end{aligned}$$

■

**定理 23.11 の証明.** 定義から

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} \rangle_k &= \sum_{I,J \in \mathcal{I}(k,N)} x^I y^J \iota_\sharp(\mathbf{u}_J)(\mathbf{u}_I) \\ &= \sum_{I,J \in \mathcal{I}(k,N)} g_{IJ} x^I y^J.\end{aligned}$$

■

**定理 23.14** $\mathbf{x}_j, \mathbf{y}_j \in X$, $j = 1, \cdots, k$ とする．このとき

$$\langle \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k, \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_k \rangle_k = \det \begin{pmatrix} \langle \mathbf{x}_1, \mathbf{y}_1 \rangle & \cdots & \langle \mathbf{x}_1, \mathbf{y}_k \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{x}_k, \mathbf{y}_1 \rangle & \cdots & \langle \mathbf{x}_k, \mathbf{y}_k \rangle \end{pmatrix}.$$

**証明.** 定理 23.10 と $\langle \cdot, \cdot \rangle_1 = \langle \cdot, \cdot \rangle$ より次の等式が得られる．

$$
\begin{aligned}
(\text{左辺}) &= \left(\iota_\sharp\left(\mathbf{y}_1\wedge\cdots\wedge\mathbf{y}_k\right)\right)\left(\mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k\right)\\
&= \langle \mathbf{y}_1\wedge\cdots\wedge\mathbf{y}_k, \iota_{\mathbf{x}_1}\wedge\cdots\wedge\iota_{\mathbf{x}_k}\rangle_{*,k}\\
&= \det\begin{pmatrix} \langle\mathbf{x}_1,\mathbf{y}_1\rangle & \cdots & \langle\mathbf{x}_1,\mathbf{y}_k\rangle\\ \vdots & \ddots & \vdots\\ \langle\mathbf{x}_k,\mathbf{y}_1\rangle & \cdots & \langle\mathbf{x}_k,\mathbf{y}_k\rangle \end{pmatrix}.
\end{aligned}
$$

■

**定義 23. 2**　$\mathbf{x}_j, \mathbf{y}_j \in X$, $j=1,\cdots,k$ とする．

$$|\mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k| = \sqrt{\langle \mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k, \mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k\rangle_k}$$

とし，これを $\mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k$ の**絶対値**という．

定理 23.14 より

$$|\mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_k| = \sqrt{\det\left(\langle\mathbf{x}_i,\mathbf{x}_j\rangle\right)} \tag{23.2}$$

である．

以下では特に $\boldsymbol{R}$ 上の線形空間として $\boldsymbol{R}^N$ の場合を考える．$1\le k\le N$ とし，$\left\{\mathbf{e}_1^{(k)},\cdots,\mathbf{e}_k^{(k)}\right\}$ を $\boldsymbol{R}^k$ の標準基底とする．

$$A = \begin{pmatrix} a_1^1 & \cdots & a_k^1\\ \vdots & \ddots & \vdots\\ a_1^N & \cdots & a_k^N\end{pmatrix} = (\mathbf{a}_1,\cdots,\mathbf{a}_k)\in\boldsymbol{R}^{N\times k}$$

とする．このとき次のことが成り立つ．

**定理 23. 15**

$$|\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_k| = \sqrt{\det A^TA}.$$

**証明．** (23.2) より

$$|\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_k| = \sqrt{\det\left(\langle\mathbf{a}_i,\mathbf{a}_j\rangle\right)} = \sqrt{\det\left(\mathbf{a}_i^T\mathbf{a}_j\right)}.$$

ここで $\det\left(\mathbf{a}_i^T\mathbf{a}_j\right)$ はグラム行列 (系 7.7) に他ならない．系 7.7 より定理が得られる．■

積分の変数変換の公式で用いられるものにヤコビアンがある．それは次の定義

に基づいてなされるものである．

**定義 23.3**　$S \in L\left(\boldsymbol{R}^M; \boldsymbol{R}^N\right)$ とし，$S$ を表現する行列を $A$ と表す．このとき

$$J_S = (J_A =) \begin{cases} \sqrt{\det\left(A^T A\right)} & M \leq N, \\ \sqrt{\det\left(A\, A^T\right)} & M > N \end{cases}$$

とおき，$S$ あるいは $A$ の**ヤコビアン**という．

ヤコビアンの計算として次のものがある．

**定理 23.16**　$S \in L\left(\boldsymbol{R}^M; \boldsymbol{R}^N\right)$ とし，$A$ を $S$ を表現する行列とする．
(1)　$M \leq N$ とする．$I \in I(M, N)$ に対して，$P_I$ を $(M, N)$ 行列で

$$P_I \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_M} \end{pmatrix}$$

をみたすものとする．このとき

$$J_S = \sqrt{\sum_{I \in I(M,N)} \left(\det\left(P_I A\right)\right)^2}$$

が成り立つ．
(2)　$M > N$ とする．$A^T$ から定まる線形写像を $S^T$ と表す．$S^T \in L\left(\boldsymbol{R}^N; \boldsymbol{R}^M\right)$ で，

$$J_S = J_{S^T}$$

が成り立つ．

**証明.** $\mathbf{a}_i = \sum_{j=1}^{N} a_i^j \mathbf{e}_j$ とすると，定理 23.1 より

$$\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k = \sum_{I \in \mathcal{I}(k,N)} \det \begin{pmatrix} a_1^{i_1} & \cdots & a_k^{i_1} \\ \vdots & \ddots & \vdots \\ a_1^{i_k} & \cdots & a_k^{i_k} \end{pmatrix} \mathbf{u}_I$$

であるが，$P_I A = \begin{pmatrix} a_1^{i_1} & \cdots & a_k^{i_1} \\ \vdots & \ddots & \vdots \\ a_1^{i_k} & \cdots & a_k^{i_k} \end{pmatrix}$ より (1) が得られる．(2) は定義から明らか．■

**注意 23.4**　測度論を知っている読者のために，$k$-ベクトルの絶対値の一つの測度論的な解釈を述べる．$1 \leq k \leq N$ とする．

$$[0,1]^k = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} : 0 \leq x_j \leq 1,\ j = 1, \cdots, k \right\}$$

とする．$S \in L\left(\boldsymbol{R}^k; \boldsymbol{R}^N\right)$ とする．$\Omega \subset \boldsymbol{R}^k$ に対して

$$S(\Omega) = \{S(\mathbf{x}) : \mathbf{x} \in \Omega\}$$

と表す．$\mathcal{H}^k$ を $\boldsymbol{R}^N$ 上の $k$ 次元ハウスドルフ測度とする．また $\mathcal{L}^n$ は ($\boldsymbol{R}^n$ 上の) $n$ 次元ルベーグ測度とする．$S$ を表現する行列を $A \in \boldsymbol{R}^{N \times k}$ とし，$\mathbf{a}_j = A(:, j)$, $j = 1, \cdots, k$ とおく．このとき，次のことが成り立つ．

$$|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k| = \mathcal{H}^k\left(S\left([0,1]^k\right)\right).$$

一般に，$k$ 次元ボレル可測集合 $\Omega \subset \boldsymbol{R}^k$ に対して，

$$\mathcal{H}^k(S(\Omega)) = J_S \mathcal{L}^k(\Omega)$$

が成り立つ．大雑把にいえば，$|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k|$ は $k$ 次元立方体 $[0,1]^k$ を $S$ で変換した図形の $k$ 次元ハウスドルフ測度であるといえる．$M > N$ で $S \in L\left(\boldsymbol{R}^M; \boldsymbol{R}^N\right)$ の場合，次の公式が成り立つことが知られている．

**定理 23.17**　$M$ 次元ボレル可測集合 $\Omega \subset \boldsymbol{R}^M$ に対して

$$\int_{\boldsymbol{R}^N} \mathcal{H}^{M-N}\left(\Omega \cap S^{-1}(\mathbf{y})\right) d\mathbf{y} = J_S \mathcal{L}^M(\Omega).$$

これを**余積公式**という．証明には測度論を必要とする．

## 23.4　*k*-形式に対する内積

$\boldsymbol{R}$ 上の $N$ 次元線形空間 $X$ に内積 $\langle\cdot,\cdot\rangle$ が定義されているとする．本節ではこの $X$ 上の内積から $\Lambda^k(X^*)$ 上の内積を定義する．

$\langle\cdot,\cdot\rangle_k$ を第 23.3 節で定義された $\Lambda^k(X)$ 上の内積とする．系 21.15 を内積空間 $\left(\Lambda^k(X), \langle\cdot,\cdot\rangle_k\right)$ に適用して得られる $\Lambda^k(X)$ から $\left(\Lambda^k(X)\right)^*$ への同形写像を $\chi_k$ と表す．定義からこのとき，任意の $\varphi \in \left(\Lambda^k(X)\right)^*$ に対して

$$\varphi(\mathbf{x}) = \left\langle \mathbf{x}, \chi_k^{-1}(\varphi) \right\rangle_k, \quad \mathbf{x} \in \Lambda^k(X)$$

が成り立つ．また $\Gamma$ を定理 23.9 で定められた $\Lambda^k(X^*)$ から $\left(\Lambda^k(X)\right)^*$ への同形写像とする．またこのとき $\mathbf{f},\mathbf{h}\in\Lambda^k(X^*)$ に対して

$$\langle\langle\mathbf{f},\mathbf{h}\rangle\rangle_k=\left\langle\chi_k^{-1}\left(\Gamma\left(\mathbf{f}\right)\right),\ \chi_k^{-1}\left(\Gamma\left(\mathbf{h}\right)\right)\right\rangle_k$$

と定義すると，明らかに $\langle\langle\cdot,\cdot\rangle\rangle_k$ は $\Lambda^k(X^*)$ の内積になっている．

特に $k=1$ の場合は $\Lambda^1(X^*)=X^*$ であり，$\Gamma(\mathbf{f})=\mathbf{f}$ であるから，

$$\langle\langle\mathbf{f},\mathbf{h}\rangle\rangle_1=\left\langle\chi_k^{-1}\left(\mathbf{f}\right),\ \chi_k^{-1}\left(\mathbf{h}\right)\right\rangle$$

となっている．

$\{\mathbf{u}_1,\cdots,\mathbf{u}_N\}$ を $X$ の基底とし，その双対基底を $\left\{\mathbf{f}^1,\cdots,\mathbf{f}^N\right\}$ とする．このとき $k$-形式は

$$\mathbf{f}=\sum_{I\in\mathcal{I}(k,N)}f_I\mathbf{f}^I,\quad \mathbf{h}=\sum_{I\in\mathcal{I}(k,N)}h_I\mathbf{f}^I$$

のように表せるが，以下では内積 $\langle\langle\mathbf{f},\mathbf{h}\rangle\rangle_k$ を係数 $f_I$, $h_I$ を用いて計算する方法を述べる．

$g_{ij}$, $g^{ij}$, $g_{IJ}$, $g^{IJ}$ は前節で定めたものとする．

**補題 23. 18**　$\mathbf{f}=\sum_{I\in\mathcal{I}(k,N)}f_I\mathbf{f}^I\in\Lambda^k(X^*)$ に対して，

$$\mathbf{a}=\sum_{I\in\mathcal{I}(k,N)}\sum_{J\in\mathcal{I}(k,N)}g^{IJ}f_J\mathbf{u}_I$$

とおく．このとき任意の $\mathbf{x}\in\Lambda^k(X)$ に対して，

$$\Gamma\left(\mathbf{f}\right)\left(\mathbf{x}\right)=\langle\mathbf{x},\mathbf{a}\rangle_k$$

である．したがって $\mathbf{a}=\chi_k^{-1}\left(\Gamma\left(\mathbf{f}\right)\right)$ である．

**証明.** $\mathbf{x}=\sum_{I\in\mathcal{I}(k,N)}x^I\mathbf{u}_I$ とする．このとき

$$\Gamma\left(\mathbf{f}\right)\left(\mathbf{x}\right)=\langle\mathbf{x},\mathbf{f}\rangle_{*,k}=\sum_{I\in\mathcal{I}(k,N)}x^If_I$$

をみたしている．一方，定理 23.11 より

$$\begin{aligned}\langle\mathbf{x},\mathbf{a}\rangle_k&=\left\langle\sum_{I\in\mathcal{I}(k,N)}x^I\mathbf{u}_I,\ \sum_{I'\in\mathcal{I}(k,N)}\sum_{J'\in\mathcal{I}(k,N)}g^{I'J'}f_{J'}\mathbf{u}_{I'}\right\rangle_k\\&=\sum_{I\in\mathcal{I}(k,N)}\sum_{I'\in\mathcal{I}(k,N)}\sum_{J'\in\mathcal{I}(k,N)}x^If_{J'}g^{I'J'}g_{II'}\end{aligned}$$

$$= \sum_{I \in \mathcal{I}(k,N)} \sum_{J' \in \mathcal{I}(k,N)} x^I f_{J'} \delta_I^{J'}$$

$$= \sum_{I \in \mathcal{I}(k,N)} x^I f_I$$

である．よって補題が証明された． ■

**定理 23.19** $\mathbf{f} = \sum_{I \in \mathcal{I}(k,N)} f_I \mathbf{f}^I$, $\mathbf{h} = \sum_{I \in \mathcal{I}(k,N)} h_I \mathbf{f}^I \in \Lambda^k (X^*)$ に対して

$$\langle\langle \mathbf{f}, \mathbf{h} \rangle\rangle_k = \sum_{I \in \mathcal{I}(k,N)} \sum_{J \in \mathcal{I}(k,N)} g^{IJ} f_I h_J$$

である．

**証明．** $\mathbf{a} = \chi_k^{-1} (\Gamma (\mathbf{f}))$, $\mathbf{b} = \chi_k^{-1} (\Gamma (\mathbf{h}))$ とおくと，補題 23.18 より

$$\mathbf{a} = \sum_{I \in \mathcal{I}(k,N)} \sum_{J \in \mathcal{I}(k,N)} g^{IJ} f_J \mathbf{u}_I, \quad \mathbf{b} = \sum_{I \in \mathcal{I}(k,N)} \sum_{J \in \mathcal{I}(k,N)} g^{IJ} h_J \mathbf{u}_I$$

となっている．したがって内積の定義から

$$\begin{aligned}
\langle\langle \mathbf{f}, \mathbf{h} \rangle\rangle_k &= \langle \mathbf{a}, \mathbf{b} \rangle_k \\
&= \sum_{I \in \mathcal{I}(k,N)} \sum_{J \in \mathcal{I}(k,N)} g_{IJ} \left( \sum_{K \in \mathcal{I}(k,N)} g^{IK} f_K \right) \left( \sum_{L \in \mathcal{I}(k,N)} g^{JL} h_L \right) \\
&= \sum_{J \in \mathcal{I}(k,N)} \sum_{L \in \mathcal{I}(k,N)} \sum_{K \in \mathcal{I}(k,N)} \left( \sum_{I \in \mathcal{I}(k,N)} g_{IJ} g^{IK} \right) f_K g^{JL} h_L \\
&= \sum_{J \in \mathcal{I}(k,N)} \sum_{L \in \mathcal{I}(k,N)} \sum_{K \in \mathcal{I}(k,N)} \delta_J^K g^{JL} f_K h_L \\
&= \sum_{L \in \mathcal{I}(k,N)} \sum_{K \in \mathcal{I}(k,N)} g^{KL} f_K h_L
\end{aligned}$$

となり，定理が証明された． ■

最後に微分形式の理論などで重要な役割を果たすホッジ $*$ 作用素の定義をする．以下では $\Lambda^k (X^*)$ を内積 $\langle\langle \cdot, \cdot \rangle\rangle_k$ による内積空間とする $(k = 1, 2, \cdots, N)$．また $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ を $X$ の正規直交基底とし，$X$ は $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ により正に向き付けられているとする．$\{\mathbf{f}^1, \cdots, \mathbf{f}^N\}$ を $\{\mathbf{u}_1, \cdots, \mathbf{u}_N\}$ の双対基底とする．このとき $\{\mathbf{f}^I : I \in I(k,N)\}$ は $\Lambda^k (X^*)$ の正規直交基底である．なぜならば $g_{ij} = \delta_{ij}$

であるので，定理 23.19 より

$$\langle\langle \mathbf{f}^I, \mathbf{f}^J \rangle\rangle_k = \begin{cases} 1, & I = J \\ 0, & I \neq J \end{cases}$$

となっているからである．

$\mathbf{f} \in \Lambda^k(X^*)$, $\mathbf{h} \in \Lambda^{N-k}(X^*)$ に対して

$$\mathbf{f} \wedge \mathbf{h} \in \Lambda^N(X^*) = \mathrm{span}\left\{\mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N\right\}$$

であるから，

$$\mathbf{f} \wedge \mathbf{h} = \alpha \mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N$$

をみたすスカラー $\alpha$ が一意的に存在する．この $\alpha$ を $F(\mathbf{f}, \mathbf{h})$ と表す．いま $\mathbf{f}$ を固定し，

$$F(\mathbf{f}, \cdot) : \Lambda^{N-k}(X^*) \ni \mathbf{h} \longmapsto F(\mathbf{f}, \mathbf{h}) \in \boldsymbol{R}$$

とすると，$F(\mathbf{f}, \cdot) \in \left(\Lambda^{N-k}(X^*)\right)^*$ となっている．ここで $\Lambda^{N-k}(X^*)$ を内積 $\langle\langle \cdot, \cdot \rangle\rangle_{N-k}$ による内積空間とし，これに定理21.15を用いれば，ある $\mathbf{q} \in \Lambda^{N-k}(X^*)$ で，任意の $\mathbf{h} \in \Lambda^{N-k}(X^*)$ に対して

$$F(\mathbf{f}, \mathbf{h}) = \langle\langle \mathbf{q}, \mathbf{h} \rangle\rangle_{N-k}$$

をみたすようなものが一意的に存在する．この $\mathbf{q}$ を ${}^*\mathbf{f}$ と表す．したがって

$$\mathbf{f} \wedge \mathbf{h} = \langle\langle {}^*\mathbf{f}, \mathbf{h} \rangle\rangle_{N-k} \mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N.$$

このとき

$${}^* : \Lambda^k(X^*) \ni \mathbf{f} \longmapsto {}^*\mathbf{f} \in \Lambda^{N-k}(X^*)$$

である．これを**ホッジ $*$ 作用素**あるいは **$*$ 作用素**という．

ホッジ $*$ 作用素の基本的な性質を証明しておく．

$I \in I(k, N)$ に対して $I^c \in I(N-k, N)$ を $I \cup I^c = \{1, \cdots, N\}$ となるようなものとする．$I \cap I^c = \varnothing$ である．$I = (i_1, \cdots, i_N) \in I(k, N)$ のとき，$I^c = (j_1, \cdots, j_{N-k})$ とし，

$$\sigma_I = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & N \\ i_1 & \cdots & i_k & j_1 & \cdots & j_{N-k} \end{pmatrix}$$

とする．

**定理 23.20**　$I = (i_1, \cdots, i_N) \in I(k, N)$ のとき，

$$^*\mathbf{f}^I = \operatorname{sgn}(\sigma_I)\ \mathbf{f}^{I^c}$$

が成り立つ．

**証明.** $\mathbf{f}^I \wedge \mathbf{h} = \langle\langle ^*\mathbf{f}^I, \mathbf{h}\rangle\rangle_{N-k}\, \mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N$ である．$J \in I(N-k, N)$ を任意にとる．$J \neq I^c$ の場合は $\mathbf{f}^I \wedge \mathbf{f}^J = \mathbf{0}$ であるから，$\langle\langle ^*\mathbf{f}^I, \mathbf{f}^J\rangle\rangle_{N-k} = 0$ であり，$J = I^c$ の場合は

$$\mathbf{f}^I \wedge \mathbf{f}^{I^c} = \left\langle\left\langle ^*\mathbf{f}^I, \mathbf{f}^{I^c}\right\rangle\right\rangle_{N-k} \mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N$$

より $\langle\langle ^*\mathbf{f}^I, \mathbf{f}^{I^c}\rangle\rangle_{N-k} = \operatorname{sgn}(\sigma_I)$ でなければならない．したがって $*$ 作用素の定義より $^*\mathbf{f}^I = \operatorname{sgn}(\sigma_I)\,\mathbf{f}^{I_c}$ が得られる． ■

**定理 23.21** (1) $\mathbf{f} \in \Lambda^k(X^*)$ に対して

$$^{**}\mathbf{f} = (-1)^{k(N-k)}\,\mathbf{f}$$

である．特に $\dim X$ が奇数ならば $^{**}\mathbf{f} = \mathbf{f}$ である．

(2) $\mathbf{f} \in \Lambda^k(X^*)$ と $\mathbf{h} \in \Lambda^{N-k}(X^*)$ に対して

$$\langle\langle ^*\mathbf{f}, \mathbf{h}\rangle\rangle_{N-k} = (-1)^{k(N-k)}\,\langle\langle \mathbf{f},\ ^*\mathbf{h}\rangle\rangle_k\,.$$

**証明.** (1) $\mathbf{f} \in \Lambda^k(X^*)$ とする．$\mathbf{f} = \sum_{I \in I(k,N)} f_I \mathbf{f}^I$ と表せている．このとき

$$^*\mathbf{f} = \sum_{I \in I(k,N)} f_I \operatorname{sgn}(\sigma_I)\ \mathbf{f}^{I^c}$$

である．また $I^{cc} = I$ であり，

$$\begin{aligned} \operatorname{sgn}(\sigma_{I^c}) &= \operatorname{sgn}\begin{pmatrix} 1 & \cdots & \cdots & \cdots & \cdots & N \\ j_1 & \cdots & j_{N-k} & i_1 & \cdots & i_k \end{pmatrix} \\ &= (-1)^{k(N-k)}\operatorname{sgn}(\sigma_I) \end{aligned}$$

であるから，

$$^*\mathbf{f}^{I^c} = (-1)^{k(N-k)}\operatorname{sgn}(\sigma_I)\,\mathbf{f}^I$$

となっている．よって (1) が証明された．

(2) 次の等式から得られる．

$$\begin{aligned} \langle\langle ^*\mathbf{f}, \mathbf{h}\rangle\rangle_{N-k}\,\mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N &= \mathbf{f} \wedge \mathbf{h} = (-1)^{k(N-k)}\,\mathbf{h} \wedge \mathbf{f} \\ &= (-1)^{k(N-k)}\,\langle\langle ^*\mathbf{h}, \mathbf{f}\rangle\rangle_k\,\mathbf{f}^1 \wedge \cdots \wedge \mathbf{f}^N. \end{aligned}$$

■

**例 23.1（ベクトル積とホッジ $*$ 作用素）**　$\boldsymbol{R}^3$ を標準基底 $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ によって正に向き付けておく．$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ の双対基底を $\{\mathbf{f}^1, \mathbf{f}^2, \mathbf{f}^3\}$ とし，これらを固定する．$\mathbf{a}, \mathbf{b} \in (\boldsymbol{R}^3)^*$ に対して，

$$\mathbf{a} \times \mathbf{b} = {}^*(\mathbf{a} \wedge \mathbf{b})$$

を $\mathbf{a}, \mathbf{b}$ の**ベクトル積**という．この演算の幾何学的な意味を述べる．$\mathbf{a} = a\mathbf{f}^1$ (ただし $a > 0$) で，$\mathbf{b} \in \mathrm{span}\{\mathbf{f}^1, \mathbf{f}^2\}$ である場合を考える．この場合，$a = \|\mathbf{a}\|$ であり，

$$\mathbf{b} = (\|\mathbf{b}\| \cos\theta)\, \mathbf{f}^1 + (\|\mathbf{b}\| \sin\theta)\, \mathbf{f}^2, \quad 0 \le \theta < 2\pi$$

と表せる．このとき

$$\begin{aligned}\mathbf{a} \wedge \mathbf{b} &= \|\mathbf{a}\|\, \mathbf{f}^1 \wedge \left\{(\|\mathbf{b}\| \cos\theta)\, \mathbf{f}^1 + (\|\mathbf{b}\| \sin\theta)\, \mathbf{f}^2\right\} \\ &= (\|\mathbf{a}\|\, \|\mathbf{b}\| \sin\theta)\, \mathbf{f}^1 \wedge \mathbf{f}^2\end{aligned}$$

が成り立つ．${}^*(\mathbf{f}^1 \wedge \mathbf{f}^2) = \mathbf{f}^3$ であるから

$$\mathbf{a} \times \mathbf{b} = (\|\mathbf{a}\|\, \|\mathbf{b}\| \sin\theta)\, \mathbf{f}^3$$

となっている．

いま空間ベクトルを座標表示 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ により，$\sum_{j=1}^{3} x_j \mathbf{f}^j \in (\boldsymbol{R}^3)^*$ とみなすことにする．このとき次のような幾何的な解釈ができる．$\mathbf{a} \times \mathbf{b}$ は $\mathbf{a}$ と $\mathbf{b}$ のなす角が 0 または $\pi$ の場合は零ベクトルである．そうでない場合は $\mathbf{a}$ から $\mathbf{b}$ への回転の向きを回転角が 0 より大きく $\pi$ 未満になるようにとると，その回転に対する右ねじの方向に $\mathbf{a} \times \mathbf{b}$ は向いていて，その長さは $\mathbf{a}$ と $\mathbf{b}$ で作られる平行四辺形の面積である．$\mathbf{a} = \sum_{j=1}^{3} a_j \mathbf{f}^j$, $\mathbf{b} = \sum_{j=1}^{3} b_j \mathbf{f}^j$ であるとき

$$\begin{aligned}\mathbf{a} \wedge \mathbf{b} &= \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_j \mathbf{f}^i \wedge \mathbf{f}^j \\ &= (a_1 b_2 - a_2 b_1)\, \mathbf{f}^1 \wedge \mathbf{f}^2 + (a_2 b_3 - a_3 b_2)\, \mathbf{f}^2 \wedge \mathbf{f}^3 + (a_3 b_1 - a_1 b_3)\, \mathbf{f}^3 \wedge \mathbf{f}^1.\end{aligned}$$

ゆえに

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2)\, \mathbf{f}^1 + (a_3 b_1 - a_1 b_3)\, \mathbf{f}^2 + (a_1 b_2 - a_2 b_1)\, \mathbf{f}^3$$

となっている．

$\mathbf{a}, \mathbf{b}, \mathbf{c} \in (\boldsymbol{R}^3)^*$ に対して $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$, $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ が定義されるが，これらは

一致するとは限らない．ベクトル積に関する公式をあげる．

**定理 23.22** $\mathbf{a}, \mathbf{b}, \mathbf{c} \in (\boldsymbol{R}^3)^*$ に対して次が成り立つ．

(1) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \langle\langle \mathbf{a}, \mathbf{c} \rangle\rangle_1 \mathbf{b} - \langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle_1 \mathbf{c}$.

(2) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$.

証明は直接計算することによってできる．

# 付録

# 付録A　置換を互換の積に分解する方法

置換が与えられたときに，それを互換に分解する一つの方法を述べておこう．まず左から最初に上の行と下の行が異なる列を見出す．たとえば

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 5 & 7 & 6 & 2 & 3 \end{pmatrix}$$

であれば第 2 列目の $\begin{matrix} 2 \\ 4 \end{matrix}$ がそれにあたる．そこで互換 $(2,4)$ を抜き出して

$$\begin{aligned} \sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 5 & 7 & 6 & 2 & 3 \end{pmatrix} \\ &= (2,4) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & \mathit{2} & 5 & 7 & 6 & \mathit{4} & 3 \end{pmatrix} \end{aligned}$$

とすることができる．実際ここで $2 \to 2 \to 4$, $6 \to 4 \to 2$ であるから，この等号が成り立つ．次に左から最初に上の行と下の行が異なる列を見出すと $\begin{matrix} 3 \\ 5 \end{matrix}$ である．そこで互換 $(3,5)$ を抜き出せば，前述の場合と同様

$$\begin{aligned} &(2,4) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & \mathit{2} & 5 & 7 & 6 & \mathit{4} & 3 \end{pmatrix} \\ &= (2,4)(3,5) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & \mathit{3} & 7 & 6 & 4 & \mathit{5} \end{pmatrix} \end{aligned}$$

となる．さらに左から最初に上の行と下の行が異なる列をみると $\begin{matrix} 4 \\ 7 \end{matrix}$ であるから，互換 $(4,7)$ を抜き出して

$$(2,4)(3,5)\begin{pmatrix}1&2&3&4&5&6&7\\1&2&\mathit{3}&7&6&4&\mathit{5}\end{pmatrix}$$
$$=(2,4)(3,5)(4,7)\begin{pmatrix}1&2&3&4&5&6&7\\1&2&3&\mathit{4}&6&\mathit{7}&5\end{pmatrix}$$

が成り立つ．この操作を繰り返していけば

$$(2,4)(3,5)(4,7)\begin{pmatrix}1&2&3&4&5&6&7\\1&2&3&\mathit{4}&6&\mathit{7}&5\end{pmatrix}$$
$$=(2,4)(3,5)(4,7)(5,6)\begin{pmatrix}1&2&3&4&5&6&7\\1&2&3&4&\mathit{5}&7&\mathit{6}\end{pmatrix}$$
$$=(2,4)(3,5)(4,7)(5,6)(6,7)$$

が得られる．すなわち

$$\begin{pmatrix}1&2&3&4&5&6&7\\1&4&5&7&6&2&3\end{pmatrix}=(2,4)(3,5)(4,7)(5,6)(6,7)$$

である．

**問題 A.1** (1) $\begin{pmatrix}1&2&3&4\\2&4&1&3\end{pmatrix}$ を互換の積に分解せよ．

(2) $\begin{pmatrix}1&2&3&4\\4&3&1&2\end{pmatrix}=(1,4)(2,3)(3,4)$ を示せ．

# 付録B　行列式の幾何学的意味

行列式はある幾何的な性質も記述している．これについて述べておこう．平面 $\boldsymbol{R}^2$ 上の基本ベクトル $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を変換 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ で写したベクトルを

$$\mathbf{u} = A\mathbf{e}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \mathbf{v} = A\mathbf{e}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

とする．$\mathbf{u} \neq \mathbf{0}, \mathbf{v} \neq \mathbf{0}$ とする．このとき次の三つの場合が考えられる．

(1)　$\mathbf{u}$ から反時計回りに $\theta\,(0 < \theta < \pi)$ 回転したところに $\mathbf{v}$ がある．

(2)　$\mathbf{u}$ から反時計回りに $\theta\,(\pi < \theta < 2\pi)$ 回転したところに $\mathbf{v}$ がある．

(3)　$\mathbf{u}$ から反時計回りに $0$ または $\pi$ 回ったところに $\mathbf{v}$ がある．

さて，

$$\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \mathbf{v} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mathbf{u} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$$

であることに注意すると

$$\begin{aligned} |A| &= a_{11}a_{22} - a_{12}a_{21} \\ &= a_{11}(a_{11}\sin\theta + a_{21}\cos\theta) - a_{21}(a_{11}\cos\theta - a_{21}\sin\theta) \\ &= a_{11}^2\sin\theta + a_{11}a_{21}\cos\theta - a_{11}a_{21}\cos\theta + a_{21}^2\sin\theta \\ &= \left(a_{11}^2 + a_{21}^2\right)\sin\theta \end{aligned}$$

である．したがって上の (1), (2), (3) のケースは，次のように行列式の符号の違いによって特徴付けられる．

(1)　$\Longleftrightarrow$　$|A| > 0$.

(2)　$\Longleftrightarrow$　$|A| < 0$.

(3) $\iff$ $|A| = 0$.

また，ベクトル $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ に対して，$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2$ とし，$|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ とする．ベクトル $\mathbf{u}, \mathbf{v}$ で作られる平行四辺形の面積を $S$ とすると，

$$\begin{aligned} S^2 &= |\mathbf{u}|^2 |\mathbf{v}|^2 \sin^2\theta = |\mathbf{u}|^2 |\mathbf{v}|^2 (1 - \cos^2\theta) \\ &= |\mathbf{u}|^2 |\mathbf{v}|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2 \\ &= \left(a_{11}^2 + a_{21}^2\right)\left(a_{12}^2 + a_{22}^2\right) - \left(a_{11}a_{12} + a_{21}a_{22}\right)^2 \\ &= \left(a_{11}a_{22} - a_{12}a_{21}\right)^2 \\ &= (\det A)^2 . \end{aligned}$$

したがって，

$$S = |\det A|$$

を得る．

# 付録C　行列に対するノルム

二つの行列がどのくらい違うかを測る尺度として，行列に対してもノルムという概念を導入する．

$(M, N)$ 行列全体からなる集合を $\boldsymbol{K}^{M\times N}$ と表す．$A, B \in \boldsymbol{K}^{M\times N}$ に対して，

$$A + B \in \boldsymbol{K}^{M\times N}.$$

また $\alpha \in \boldsymbol{K}$ に対して，

$$\alpha A \in \boldsymbol{K}^{M\times N}$$

である．行列に対するノルムは次のように定義される．

**定義 C.1**　$\boldsymbol{K}^{M\times N}$ から $\boldsymbol{R}$ への関数 $\rho(\cdot)$ が次の条件 (1), (2), (3), (4) をみたすとき，$\boldsymbol{K}^{M\times N}$ 上の**ノルム**であるという．

(1)　$\rho(A) \geq 0, \quad A \in \boldsymbol{K}^{M\times N}.$

(2)　$\rho(A) = 0 \Longleftrightarrow A = O.$

(3)　$\rho(\alpha A) = |\alpha|\,\rho(A), \quad \alpha \in \boldsymbol{K},\ A \in \boldsymbol{K}^{M\times N}.$

(4)　$\rho(A+B) \leq \rho(A) + \rho(B), \quad A, B \in \boldsymbol{K}^{M\times N}.$

$\boldsymbol{K}^{M\times N}$ 上のノルムの例としては本文でも解説した作用素ノルムがある．ここではそれ以外のノルムについて述べる．

**定義 C.2**　$A = (a_{jk}) \in \boldsymbol{K}^{M\times N}$ に対して，

$$\|A\|_F = \left(\sum_{j=1}^{M}\sum_{k=1}^{N} |a_{jk}|^2\right)^{\frac{1}{2}}$$

を**フロベニウス・ノルム**あるいは**ヒルベルト-シュミット・ノルム**という．

フロベニウス・ノルムが $\boldsymbol{K}^{M\times N}$ 上のノルムであることは，$A = (a_{jk})$ を縦

に $a_{11}, \cdots, a_{1n}, a_{21}, \cdots, a_{2n}, \cdots, a_{m1}, \cdots, a_{mn}$ の順に並べて $\boldsymbol{K}^{MN}$ の要素と考えれば，$\|A\|_F$ は $\boldsymbol{K}^{MN}$ 上の $l^2$ ノルムに他ならないことから証明できる．

このノルムは次の性質をもっている．

**命題 C. 1** (1) $A \in \boldsymbol{K}^{M\times N}$ ならば

$$\|A\|_F^2 = \mathrm{tr}(A^*A).$$

(2) $A \in \boldsymbol{K}^{M\times N}$, $\mathbf{x} \in \boldsymbol{K}^N$ に対して，

$$\|A\mathbf{x}\|_2 \le \|A\|_F \|\mathbf{x}\|_2 .$$

(3) $A, B \in \boldsymbol{K}^{N\times N}$ ならば

$$\|AB\|_F \le \|A\|_F \|B\|_F .$$

**証明.** $(A^*A)(j,k) = \sum_{l=1}^{M} \overline{A(l,j)}\, A(l,k)$ より

$$\mathrm{tr}(A^*A) = \sum_{k=1}^{N}\sum_{l=1}^{M} \overline{A(l,k)}\, A(l,k) = \|A\|_F^2$$

である．また，$(A\mathbf{x})(j) = \sum_{k=1}^{N} A(j,k)\, x_k = \langle A(j,:), \mathbf{x}\rangle$ より

$$\|A\mathbf{x}\|_2^2 = \sum_{j=1}^{M} |(A(j,:), \mathbf{x})|^2 \le \sum_{j=1}^{M} \|A(j,:)\|_2^2 \|\mathbf{x}\|_2^2 = \|A\|_F^2 \|\mathbf{x}\|_2^2$$

となっている．$\mathbf{b}_j = B(;,j)$ とおく．

$$AB = A\begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_N \end{pmatrix} = \begin{pmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_N \end{pmatrix}$$

であるから，

$$\|AB\|_F^2 = \sum_{k=1}^{N} \|A\mathbf{b}_k\|_2^2 \le \sum_{k=1}^{N} \|A\|_F^2 \|B(;,k)\|_2^2 = \|A\|_F^2 \|B\|_F^2 .$$

■

$\|\cdot\|_{op}$ の考え方を一般化した行列ノルムを紹介しておこう．なお『有界閉集合上の連続関数は，有限な最大値と最小値をもつ』という定理を用いる．

**定理 C. 2** $\rho(\cdot)$ を $\boldsymbol{K}^M$ のノルム，$\mu(\cdot)$ を $\boldsymbol{K}^N$ のノルムとする．このとき $A \in \boldsymbol{K}^{M\times N}$ に対して，

$$\|A\|_{\mu,\rho} = \sup\left\{\rho(A\mathbf{x}) : \mathbf{x} \in \boldsymbol{K}^N,\ \mu(\mathbf{x}) = 1\right\}$$

は $\boldsymbol{K}^{M\times N}$ 上のノルムである．

**証明.** 問題 10.1 よりノルムは連続関数であるから，$B=\left\{\mathbf{x}\in\boldsymbol{K}^N:\mu(\mathbf{x})=1\right\}$ は $\boldsymbol{K}^N$ の閉集合である．また問題 10.2 より $B$ は有界集合である．$\mathbf{x}\mapsto A\mathbf{x}$ は明らかに連続写像であるから，$\mathbf{x}\mapsto\rho(A\mathbf{x})$ は連続関数である．ゆえにこの連続関数は $B$ 上で最大値をとる．したがって $\|A\|_{\mu,\rho}$ は定義される．もし $\|A\|_{\mu,\rho}=0$ ならば，任意の $\mathbf{x}\in B$ に対して，$A\mathbf{x}=\mathbf{0}$ である．$\mu(\mathbf{e}_j)^{-1}\mathbf{e}_j\in B$ であり，

$$\mathbf{0}=A\mu(\mathbf{e}_j)^{-1}\mathbf{e}_j=\mu(\mathbf{e}_j)^{-1}A\mathbf{e}_j$$

より $A\mathbf{e}_j=\mathbf{0}$ である．ゆえに $A=O$ となる．行列ノルムの条件の (3) は明らかである．

$$\rho((A+B)\mathbf{x})\leq\rho(A\mathbf{x})+\rho(B\mathbf{x})$$

より，行列ノルムの条件の (4) も明らかに成り立つ．■

**命題 C.3**　定理 C.2 の $\boldsymbol{K}^{M\times N}$ 上のノルムは次をみたす．

(1)　$A\in\boldsymbol{K}^{M\times N}$, $\mathbf{x}\in\boldsymbol{K}^N$ に対して，

$$\rho(A\mathbf{x})\leq\|A\|_{\mu,\rho}\,\mu(\mathbf{x}).$$

(2)　$M=N$, $\rho=\mu$ の場合，$A,B\in\boldsymbol{K}^{N\times N}$ に対して，

$$\|AB\|_{\rho,\rho}\leq\|A\|_{\rho,\rho}\|B\|_{\rho,\rho}.$$

**証明.** (1)　$\mathbf{x}=\mathbf{0}$ ならばこの不等式は明らかである．
$\mathbf{x}\neq\mathbf{0}$ の場合，$\mu\left(\mu(\mathbf{x})^{-1}\mathbf{x}\right)=1$ であるから，

$$\|A\|_{\mu,\rho}\geq\rho\left(A\mu(\mathbf{x})^{-1}\mathbf{x}\right)=\mu(\mathbf{x})^{-1}\rho(A\mathbf{x})$$

より命題が証明される．

(2)　(1) より

$$\rho(AB\mathbf{x})\leq\|A\|_{\rho,\rho}\,\rho(B\mathbf{x})\leq\|A\|_{\rho,\rho}\|B\|_{\rho,\rho}\,\rho(\mathbf{x})$$

が成り立つ．よって (2) が証明された．■

定理 C.2 の特別な場合として次のノルムがある．

**例 C.1**（$(p,q)$ ノルム）　$1\leq p,\,q\leq\infty$ とする．$A=(a_{jk})\in\boldsymbol{K}^{M\times N}$ に対して，

$$\|A\|_{p,q}=\sup\left\{\|A\mathbf{x}\|_q:\mathbf{x}\in\boldsymbol{K}^N,\ \|\mathbf{x}\|_p=1\right\}$$

を $(p, q)$ ノルムという．これは $\boldsymbol{K}^{M\times N}$ 上のノルムである．特に $(2,2)$ ノルムが $\|\cdot\|_{op}$ と一致している．

# 付録D　ジョルダン標準形

定理 13.15 で正方行列が対角化できるための必要十分条件を述べた．対角化できるとは限らない正方行列に対して，対角化に近い単純な形に変形する方法がジョルダン標準化である．本付録ではジョルダン標準化の具体的な計算方法の概略を述べておこう．

次の形の正方行列

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

を**ジョルダン・ブロック**という．次の定理が成り立つ．

**定理 D.1（ジョルダン標準化）**　$A$ を $N$ 次正方行列とする．ある $N$ 次正則行列 $P$ により

$$P^{-1}AP = \begin{pmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_n \end{pmatrix}$$

(ここで $J_1, \cdots, J_n$ はジョルダン・ブロック) とすることができる．右辺を $A$ の**ジョルダン標準形**という．

ジョルダン標準化を与える正則行列 $P$ は次のような手順で求めることができる．

$\lambda_1, \cdots, \lambda_s$ を $A$ の相異なる固有値とする．$\Delta_A(z)$ を $A$ の特性多項式とし，$z^m$ の部分に形式的に $A^m$ を代入すると，

$$\Delta_A(A) = O$$

が成り立つことが知られている (**ケーリー-ハミルトンの定理**)．多項式 $\psi(z)$ で $\psi(A) = O$ をみたす次数最小のものを**最小多項式**という．最小多項式は特性多項式を割り切るので，

$$\psi(z) = (z - \lambda_1)^{m_1} \cdots (z - \lambda_s)^{m_s}$$

とおくことができる．$(A - \lambda I_N)^k \mathbf{u} = \mathbf{0}$ がある正の整数 $k$ に対して成り立つベクトル $\mathbf{u}$ を $\lambda$ に属する**一般化固有ベクトル**という．特に

$$(A - \lambda I_N)^k \mathbf{u} = \mathbf{0}, \quad (A - \lambda I_N)^{k-1} \mathbf{u} \neq \mathbf{0}$$

をみたすときは**高さ $k$ の一般化固有ベクトル**という．

いま $\Omega_j = \left\{\mathbf{u} \in \boldsymbol{K}^N : (A - \lambda_j I_N)^{m_j} \mathbf{u} = \mathbf{0}\right\}$ とおく．$\Omega_j$ は $\boldsymbol{K}^N$ の部分空間になっていて，

$$\boldsymbol{K}^N = \Omega_1 \dotplus \cdots \dotplus \Omega_s,$$
$$\dim \Omega_j = \lambda_j \text{の重複度}$$

が成り立つ (証明は略す)．

まず $\Omega_1$ について述べる．$m = m_1$ とおく．$m \geq 2$ の場合を考える．$\lambda_j$ の一般固有ベクトルの高さは $m_j$ より大きくはない．

$$M_k = \left\{\mathbf{u} \in \boldsymbol{K}^N : (A - \lambda_1 I_N)^k \mathbf{u} = \mathbf{0}\right\}$$

とすると

$$\Omega_1 = M_m \supset M_{m-1} \supset \cdots \supset M_1 \supset M_0 = \{\mathbf{0}\}$$

となっている．そこで部分空間 $N_k$ を

$$M_m = N_m \dotplus M_{m-1},$$
$$M_{m-1} = N_{m-1} \dotplus M_{m-2},$$
$$\vdots$$
$$M_1 = N_1 \dotplus M_0$$

となるようにとり固定する．以下

$$A_{\lambda_1} = A - \lambda_1 I_N$$

と略記する．

$N_m$ の基底をとり，$\left\{\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m}$ とおく $(r_m = \dim N_m)$．このとき，$A_{\lambda_1}\mathbf{b}_1^{(m)}$, $\cdots, A_{\lambda_1}\mathbf{b}_{r_m}^{(m)}$ は高さ $m-1$ の一般化固有ベクトルであり，一次独立になっている．したがって

$$\left\{A_{\lambda_1}\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m} \cup \left\{\mathbf{b}_j^{(m-1)}\right\}_{j=1}^{r_{m-1}}$$

が $N_{m-1}$ の基底となるように $\mathbf{b}_1^{(m-1)}, \cdots, \mathbf{b}_{r_{m-1}}^{(m-1)} \in M_{m-1}$ をとることができる．$\dim N_{m-1} = r_m + r_{m-1}$ である．このとき，

$$A_{\lambda_1}^2\mathbf{b}_1^{(m)}, \quad \cdots, \quad A_{\lambda_1}^2\mathbf{b}_{r_m}^{(m)}, \quad A_{\lambda_1}\mathbf{b}_1^{(m-1)}, \quad \cdots, \quad A_{\lambda_1}\mathbf{b}_{r_{m-1}}^{(m-1)}$$

は高さ $m-2$ の一般化固有ベクトルで，一次独立になっている．したがって，

$$\left\{A_{\lambda_1}^2\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m} \cup \left\{A_{\lambda_1}\mathbf{b}_j^{(m-1)}\right\}_{j=1}^{r_{m-1}} \cup \left\{\mathbf{b}_j^{(m-2)}\right\}_{j=1}^{r_{m-2}}$$

が $N_{m-2}$ の基底になるように $\mathbf{b}_1^{(m-2)}, \cdots, \mathbf{b}_{r_{m-2}}^{(m-2)} \in M_{m-2}$ をとることができる．$\dim N_{m-2} = r_m + r_{m-1} + r_{m-2}$ である．以下この操作を続けることにより，

$$\left\{A_{\lambda_1}^{m-1}\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m} \cup \cdots \cup \left\{A_{\lambda_1}\mathbf{b}_j^{(2)}\right\}_{j=1}^{r_2} \cup \left\{\mathbf{b}_j^{(1)}\right\}_{j=1}^{r_1}$$

が $N_1$ の基底になるように $\mathbf{b}_1^{(1)}, \cdots, \mathbf{b}_{r_1}^{(1)} \in M_1$ を選ぶことができる．これらを

$$\left\{\mathbf{c}_{j,1}^{(m)}\right\}_{j=1}^{r_m} = \left\{\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m}$$

$$\left\{\mathbf{c}_{j,2}^{(m)}\right\}_{j=1}^{r_m} = \left\{A_{\lambda_1}\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m}, \quad \left\{\mathbf{c}_{j,1}^{(m-1)}\right\}_{j=1}^{r_{m-1}} = \left\{\mathbf{b}_j^{(m-1)}\right\}_{j=1}^{r_{m-1}}$$

$$\left\{\mathbf{c}_{j,3}^{(m)}\right\}_{j=1}^{r_m} = \left\{A_{\lambda_1}^2\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m}, \quad \left\{\mathbf{c}_{j,2}^{(m-1)}\right\}_{j=1}^{r_{m-1}} = \left\{A_{\lambda_1}\mathbf{b}_j^{(m-1)}\right\}_{j=1}^{r_{m-1}},$$

$$\left\{\mathbf{c}_{j,1}^{(m-2)}\right\}_{j=1}^{r_{m-2}} = \left\{\mathbf{b}_j^{(m-2)}\right\}_{j=1}^{r_{m-2}}$$

$$\vdots$$

$$\left\{\mathbf{c}_{j,m}^{(m)}\right\}_{j=1}^{r_m} = \left\{A_{\lambda_1}^m\mathbf{b}_j^{(m)}\right\}_{j=1}^{r_m}, \quad \cdots, \quad \left\{\mathbf{c}_{j,2}^{(2)}\right\}_{j=1}^{r_2} = \left\{A_{\lambda_1}\mathbf{b}_j^{(2)}\right\}_{j=1}^{r_2},$$

$$\left\{\mathbf{c}_{j,1}^{(1)}\right\}_{j=1}^{r_1} = \left\{\mathbf{b}_j^{(1)}\right\}_{j=1}^{r_1}$$

と表す．

$$S_j^{(m)} = \begin{pmatrix} \mathbf{c}_{j,m}^{(m)} & \mathbf{c}_{j,m-1}^{(m)} & \cdots & \mathbf{c}_{j,1}^{(m)} \end{pmatrix}, \quad j = 1, \cdots, r_m$$

$$S_j^{(m-1)} = \begin{pmatrix} \mathbf{c}_{j,m-1}^{(m-1)} & \mathbf{c}_{j,m-2}^{(m-1)} & \cdots & \mathbf{c}_{j,1}^{(m-1)} \end{pmatrix}, \quad j = 1, \cdots, r_{m-1}$$

$$\vdots$$

$$S_j^{(1)} = \begin{pmatrix} \mathbf{c}_{j,1}^{(1)} \end{pmatrix}, \quad j = 1, \cdots, r_1$$

とおき，

$$S_1 = \begin{pmatrix} S_1^{(m)} & \cdots & S_{r_m}^{(m)} & \cdots & S_1^{(1)} & \cdots & S_{r_1}^{(1)} \end{pmatrix}$$

とおく．以下 $\Omega_2, \cdots, \Omega_s$ に対して同様の考察をして，行列

$$S_2, \quad \cdots, \quad S_s$$

を求めることができる．このとき，

$$P = \begin{pmatrix} S_1 & \cdots & S_s \end{pmatrix}$$

とすると，これがジョルダン標準化を行う正則行列の一つである．

**例 D. 1** $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ のジョルダン標準化を行う．$A$ の固有値は 1 で重複度は 3 である．また $A$ の 1 に関する固有ベクトルは

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

であるから，対角化はできない．$(A - I_3)^2 = O$ である．$M_2 = \boldsymbol{K}^3$ である．$M_2 = N_2 \dotplus M_1$ である．そこで高さ 2 の一般化固有ベクトル $\mathbf{b}_1^{(2)}$ を求める．すなわち $(A - I_3)^2\,\mathbf{b}_1^{(2)} = \mathbf{0}$, $(A - I_3)\,\mathbf{b}_1^{(2)} \neq \mathbf{0}$ なる $\mathbf{b}_1^{(2)}$ を求める．たとえば $\mathbf{b}_1^{(2)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ がある．$\mathbf{c}_{1,1}^{(2)} = \mathbf{b}_1^{(2)}$ とおく．

$$\mathbf{c}_{1,2}^{(2)} = (A - I_3)\,\mathbf{b}_1^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

である．これは高さ 1 の固有ベクトルである．また $\mathbf{c}_{1,1}^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ も高さ 1 の一般化固有ベクトルである．ゆえに

$$P = \left(\mathbf{c}_{1,2}^{(2)},\ \mathbf{c}_{1,1}^{(2)},\ \mathbf{c}_{1,1}^{(1)}\right) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

でジョルダン標準化を求めることができる．実際，

$$P^{-1}AP = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

ジョルダン標準形の詳細は，たとえば［韓他］にある．

# 付録E　問題の解答

**問題 2.1**　(1)　$AB = (ac + bd)\,, BA = \begin{pmatrix} ac & bc \\ ad & bd \end{pmatrix}$．　(2)　$AB = \begin{pmatrix} 1 & 5 & 4 \\ 2 & 7 & 6 \\ 2 & 11 & 8 \end{pmatrix}$，
$BA = \begin{pmatrix} 3 & 7 & 9 \\ 2 & 3 & 4 \\ 4 & 8 & 10 \end{pmatrix}$．　(3)　$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$，$BA = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & -2 & 0 \end{pmatrix}$．

**問題 2.2**　$\mathbf{x} = (x_j)$ とすると，$\mathbf{x}^*\mathbf{x} = \sum_{j=1}^{N} \overline{x_j} x_j = \sum_{j=1}^{N} |x_j|^2 \geq 0$ である．また，これより後半の主張は明らか．

**問題 2.3**　(1)　$\begin{pmatrix} ax & 0 \\ bx + cy & cz \end{pmatrix}$　(2)　$(a_{ij})$ が下三角行列とは，$a_{ij} = 0\,(i < j)$ であることに注意する．$A = (a_{ij})\,, B = (b_{ij})$ が下三角行列であるとする．$AB$ の $(i, j)$ 成分を $c_{ij}$ とする．$c_{ij} = \sum_{k=1}^{N} a_{ik} b_{kj} = \sum_{k=1}^{i} a_{ik} b_{kj}$ である．最後の和において $i < j$ ならば $k \leq i < j$ より $b_{kj} = 0$ となるので，$c_{ij} = 0$ である．ゆえに $AB$ は下三角行列である．また $c_{ii} = \sum_{k=1}^{i} a_{ik} b_{ki} = a_{ii} b_{ii}$ である．上三角行列の場合も同様に示せる．

**問題 3.1**　$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ とすると，

$$AB = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

であるから，

$$T \circ S(\mathbf{x}) = AB\mathbf{x} = \begin{pmatrix} (aa' + bc')\,x_1 + (ca' + dc')\,x_2 \\ (ab' + bd')\,x_1 + (cb' + dd')\,x_2 \end{pmatrix}.$$

**問題 4.1**

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & a_{N3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & \lambda a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{N1} & \lambda a_{N2} & a_{N3} \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & a_{N3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \lambda a_{12} + a_{13} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & \lambda a_{N2} + a_{N3} \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & a_{N3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N3} & a_{N2} \end{pmatrix}.$$

一般に，右からかける場合は

$$\begin{cases} E_i^{(N)}(\lambda) & \cdots \quad \text{第}\,i\,\text{列を}\,\lambda\,\text{倍する.} \\ E_{i,j}^{(N)}(\lambda) & \cdots \quad \text{第}\,j\,\text{列に第}\,i\,\text{列の}\,\lambda\,\text{倍を加える.} \\ F_{i,j}^{(N)} & \cdots \quad \text{第}\,i\,\text{列と第}\,j\,\text{列を入れ替える.} \end{cases}$$

となっている．$(N, M)$ 行列 $A$ に，$N$ 次基本行列の有限個の積を右側からかけることによって得られる行列を $A$ の**列に関する基本変形**あるいは**列に関する初等変形**という．

**問題 4.2**　$\gamma = \dfrac{\alpha}{\alpha \mathbf{y}^* \mathbf{x} - 1}$ であるから，$E_i(\lambda)$ については

$$\gamma = \frac{1-\lambda}{(1-\lambda)-1} = 1 - \lambda^{-1}$$

より

$$E_i(\lambda)^{-1} = I_N - \gamma\, \mathbf{e}_i \mathbf{e}_i^T = I_N - \left(1 - \lambda^{-1}\right) \mathbf{e}_i \mathbf{e}_i^T = E_i\left(\lambda^{-1}\right)$$

である．$E_{i,j}(\lambda)$ については $\gamma = \dfrac{-\lambda}{-1} = \lambda$ より $E_{i,j}(\lambda)^{-1} = I_N - \lambda\, \mathbf{e}_i \mathbf{e}_i^T = E_{i,j}(-\lambda)$ となっている．$F_{i,j} = I_N - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T$ については $(\mathbf{e}_i - \mathbf{e}_j)^T (\mathbf{e}_i - \mathbf{e}_j) = 2$ であるから，$\gamma = 1$ より $F_{i,j}^{-1} = F_{i,j}$ である．

**問題 4.3**　(1)

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

(2)

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & -2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 2 & -2 & 1 & 1 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right).$$

**問題 4.5** (1)

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -13 & -5 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & \frac{5}{13} & -\frac{1}{13} \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -\frac{2}{13} & \frac{3}{13} \\ 0 & 1 & \frac{5}{13} & -\frac{1}{13} \end{array}\right).$$

(2)

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 & 1 & 0 \\ 0 & -4 & 1 & -1 & -2 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 & 1 & 0 \\ 0 & -4 & 1 & -1 & -2 & 0 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 3 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & -7 & 3 & 2 & -4 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & -7 & 3 & 2 & -4 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{16}{3} & -\frac{1}{3} & -\frac{5}{3} & \frac{7}{3} & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{16} & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & \frac{17}{16} & \frac{5}{16} & -\frac{7}{16} & -\frac{3}{16} \\ 0 & 1 & -2 & 0 & \frac{17}{16} & -\frac{11}{16} & -\frac{7}{16} & -\frac{3}{16} \\ 0 & 0 & 1 & 0 & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} & -\frac{1}{16} \\ 0 & 0 & 0 & 1 & -\frac{1}{16} & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & \frac{17}{16} & \frac{5}{16} & -\frac{7}{16} & -\frac{3}{16} \\ 0 & 1 & 0 & 0 & \frac{7}{16} & \frac{3}{16} & -\frac{1}{16} & -\frac{5}{16} \\ 0 & 0 & 1 & 0 & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} & -\frac{1}{16} \\ 0 & 0 & 0 & 1 & -\frac{1}{16} & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{3}{16} & -\frac{1}{16} & -\frac{5}{16} & \frac{7}{16} \\ 0 & 1 & 0 & 0 & \frac{7}{16} & \frac{3}{16} & -\frac{1}{16} & -\frac{5}{16} \\ 0 & 0 & 1 & 0 & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} & -\frac{1}{16} \\ 0 & 0 & 0 & 1 & -\frac{1}{16} & -\frac{5}{16} & \frac{7}{16} & \frac{3}{16} \end{array}\right).$$

**問題 5.1** $\tau\sigma = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$,

$$\sigma\tau = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

$$\sigma\tau\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

$$\tau\sigma\tau^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

**問題 5.2**

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & \cdots & j & \cdots & i & \cdots & N \\ 1 & 2 & \cdots & i & \cdots & j & \cdots & N \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & N \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & N \end{pmatrix} = \sigma.$$

**問題 5.3**

$$\begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} \underset{(3\text{ 行}-1\text{ 行})}{=} \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & -2 & 1 \end{vmatrix} \underset{(3\text{ 行}+2\times(2\text{ 行}))}{=} \begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{vmatrix} \underset{\text{例題 }5.2(2)}{=} 5.$$

**問題 5.4**　基本行列は正則であるから，系 5.10 よりその行列式は 0 ではない．したがって第 4.2.1 節で述べたガウスの消去法において，変形の途中で第 1 列から第 $N$ 列の中のある列の成分がすべて 0 になるならば $|A| = 0$ となり，$A$ の正則性に反する．もしそのようなことがなければ，第 4.2.1 節のガウスの記号を用いれば $a_{11}^{(1)} a_{22}^{(2)} \cdots a_{NN}^{(N)} \neq 0$ となるように選べる．

**問題 6.1**　$A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ とすると，

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -1 & -5 & 8 \\ 1 & 1 & -2 \\ 1 & 7 & -10 \end{pmatrix}$$

であるから，

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 & -5 & 8 \\ 1 & 1 & -2 \\ 1 & 7 & -10 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ \frac{1}{2} \\ \frac{5}{2} \end{pmatrix}.$$

**問題 6.2**

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & 1 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad |A| = 2$$

より，

$$x_1 = \frac{1}{|A|}\begin{vmatrix} 1 & 3 & 1 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 0 & 2 & -1 \\ 0 & -1 & -1 \\ 1 & 1 & 2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 2 & -1 \\ -1 & -1 \end{vmatrix} = -\frac{3}{2},$$

$$x_2 = \frac{1}{|A|}\begin{vmatrix} 2 & 1 & 1 \\ 4 & 2 & 3 \\ 3 & 1 & 2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 2 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{vmatrix} = -\frac{1}{2}\begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = \frac{1}{2},$$

$$x_3 = \frac{1}{|A|}\begin{vmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 1 & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 2 & 3 & 1 \\ 0 & -5 & 0 \\ 1 & -2 & 0 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 0 & -5 \\ 1 & -2 \end{vmatrix} = \frac{5}{2}.$$

**問題 8.1**　$A_{11} = 1$, $A_{12} = (0 \quad 1)$, $A_{21} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $A_{22} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ とブロック分けする．$A_{21}$ は零行列であるから，$A_{11} - A_{12}A_{22}^{-1}A_{21} = A_{11} = 1$ で，$B_{11} = 1$,

$$B_{12} = -\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \end{pmatrix},$$

$$B_{21} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B_{22} = A_{22}^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}.$$

したがって，$A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

**問題 9.1**　$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{a}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ とする．明らかに，この中の任意の二つのベクトルは線形独立であるが，$\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_3 = \mathbf{0}$ である．

**問題 10.1**　$0 < \rho(\mathbf{e}_j) < \infty$ であるから，$c = \min_{1 \le j \le N} \rho(\mathbf{e}_j)$, $C = \max_{1 \le j \le N} \rho(\mathbf{e}_j)$ とおく．このとき，

$$\rho(\mathbf{x}) = \rho\left(\sum_{j=1}^{N} x_j \mathbf{e}_j\right) \le \sum_{j=1}^{N} |x_j|\,\rho(\mathbf{e}_j) \le C\sum_{j=1}^{N} |x_j| \le C\sqrt{N}\,\|\mathbf{x}\|$$

が成り立っている．したがって，$\delta = \dfrac{\varepsilon}{C\sqrt{N}}$ とおくと，

$$|\rho(\mathbf{x}+\mathbf{h}) - \rho(\mathbf{x})| \le \rho(\mathbf{h}) \le C\sqrt{N}\,\|\mathbf{h}\| < \varepsilon.$$

**問題 10.2**　$B = \left\{\mathbf{x} \in \boldsymbol{K}^N : \|\mathbf{x}\| = 1\right\}$ とする．このとき，定義より $\mathbf{x} \in B$ ならば $\rho(\mathbf{x}) > 0$ である．また $\rho$ は連続関数で，$B$ は有界閉集合であるから，$\rho$ は $B$ 上のある点 $\mathbf{x}'$ で最小値 $c$ をとる．$c = \rho(\mathbf{x}') > 0$ である．いま，任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{x} \neq \mathbf{0}$ に対して，$\|\mathbf{x}\|^{-1}\mathbf{x} \in B$ であるから，

$$c \leq \rho\left(\|\mathbf{x}\|^{-1}\mathbf{x}\right) = \|\mathbf{x}\|^{-1}\rho(\mathbf{x})$$

より，左側の不等式が証明できる．$\mathbf{x} = \mathbf{0}$ の場合は両辺 0 に等しく，不等式は成り立つ．右側の不等式は前問題の解答の中で証明した．

**問題 10.3**　$(PP)(j,k) = \sum_{m=1}^{N} P(j,m)\,P(m,k) = \dfrac{1}{N} = P(j,k)$ であるから，$P$ は射影行列，また明らかにエルミート行列である．

$$QQ = (I_N - P)(I_N - P) = I_N - 2P + P^2 = I_N - P = Q$$

であり，また明らかに $Q$ はエルミート行列である．$m = N^{-1}\sum_{j=1}^{N} x_j$ とおく．$P\mathbf{x} = (m \quad \cdots \quad m)^T$ である．ゆえに

$$\mathbf{z} = Q\mathbf{x} = \begin{pmatrix} x_1 - m & \cdots & x_N - m \end{pmatrix}^T$$

であり，$\sum_{j=1}^{N} z_j = \sum_{j=1}^{N}(x_j - m) = 0$ となる．逆に $\mathbf{z}$ が $\sum_{j=1}^{N} z_j = 0$ をみたすならば，$P\mathbf{z} = \mathbf{0}$ より $Q\mathbf{z} = \mathbf{z}$ である．よって問題の最後の主張が証明された．

**問題 10.4**　$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ とする．このとき，$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ x \end{pmatrix}$ であるから，

$$\mathcal{R}(A) = \left\{\begin{pmatrix} 0 \\ x \end{pmatrix} : x \in \boldsymbol{K}^N\right\},$$

$$\mathcal{N}(A) = \left\{\begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \boldsymbol{K}^N\right\}$$

である．したがって $\mathcal{R}(A) = \mathcal{N}(A)$ なので，$\mathcal{R}(A) \dotplus \mathcal{N}(A) \neq \boldsymbol{K}^N$.

**問題 10.5**　問題 10.3 の記号を用いる．明らかに $\operatorname{rank}(P) = 1$ である．$\boldsymbol{K}^N = \mathcal{R}(Q) \oplus \mathcal{R}(P)$ であるから，$\operatorname{rank}(Q) = \dim \mathcal{R}(Q) = N - \dim \mathcal{R}(P) = N - 1$ である．

**問題 10.6**　系 10.29 より，$\boldsymbol{K}^M = \mathcal{R}(A^*) \oplus \mathcal{N}(A)$ である．したがって，任意の $\mathbf{x} \in \boldsymbol{K}^M$ に対して，$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, $\mathbf{x}_1 \in \mathcal{R}(A^*)$, $\mathbf{x}_2 \in \mathcal{N}(A)$ が成り立つ．ゆえに

$$T_A(\mathbf{x}) = T_A(\mathbf{x}_1) + T_A(\mathbf{x}_2) = T_A(\mathbf{x}_1)$$

であるから，$\mathcal{R}(T_A) \subset \mathcal{R}(T_A|_{\mathcal{R}(A^*)}) \subset \mathcal{R}(T_A)$ であることがわかる．ゆえに $\mathcal{R}(T_A|_{\mathcal{R}(A^*)}) = \mathcal{R}(T_A) = \mathcal{R}(A)$ である．したがって $T_A|_{\mathcal{R}(A^*)}$ は $\mathcal{R}(A^*)$ から $\mathcal{R}(A)$ への全射である．

次に単射であることを示す．$\mathbf{x}, \mathbf{x}' \in \mathcal{R}(A^*)$ かつ $T_A|_{\mathcal{R}(A^*)}(\mathbf{x}) = T_A|_{\mathcal{R}(A^*)}(\mathbf{x}')$ をみたすとする．このとき，$T_A(\mathbf{x}) = T_A(\mathbf{x}')$ でもあるから $\mathbf{x} - \mathbf{x}' \in \mathcal{N}(A)$ となっている．したがって，

$$\mathbf{x} - \mathbf{x}' \in \mathcal{N}(A) \cap \mathcal{R}(A^*) = \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{\mathbf{0}\}$$

より $\mathbf{x} = \mathbf{x}'$ であることがわかる．以上のことから，写像 $T_A|_{\mathcal{R}(A^*)} : \mathcal{R}(A^*) \to \mathcal{R}(A)$ の逆写像 $(T_A|_{\mathcal{R}(A^*)})^{-1} : \mathcal{R}(A) \to \mathcal{R}(A^*)$ が存在し，$(T_A|_{\mathcal{R}(A^*)})^{-1}$ は $\mathcal{R}(A)$ から $\mathcal{R}(A^*)$ への全単射であることがわかる．$(T_A|_{\mathcal{R}(A^*)})^{-1}$ が線形写像であることは定理 3.4 と同様の証明により示される．

**問題 10.7**　$P$ を $\boldsymbol{K}^L$ から $\mathcal{R}(B^*)$ への直交射影とする．$T_B$ を $B$ から定まる線形写像とする．このとき，$T_B P$ は $\boldsymbol{K}^L$ から $R(B)$ への線形写像になっている．また $\mathcal{R}(A) = \mathcal{R}(B)$ であるから，問題 10.6 より $(T_A|_{\mathcal{R}(A^*)})^{-1}$ は $\mathcal{R}(B)$ から $\mathcal{R}(A^*)$ への線形写像になっている．したがって，$(T_A|_{\mathcal{R}(A^*)})^{-1} T_B P$ は $\boldsymbol{K}^L$ から $\mathcal{R}(A^*)$ への線形写像である．$\mathbf{x} \in \boldsymbol{K}^L$ に対して，$\mathbf{x} = \mathbf{x}' + \mathbf{x}''$, $\mathbf{x}' \in \mathcal{R}(B^*)$, $\mathbf{x}'' \in \mathcal{R}(B^*)^\perp = \mathcal{N}(B)$ と一意的に分解される．$T_B(\mathbf{x}') \in \mathcal{R}(B) = \mathcal{R}(A)$ であるから，

$$T_A\,(T_A|_{\mathcal{R}(A^*)})^{-1}\,T_B(\mathbf{x}') = (T_A|_{\mathcal{R}(A^*)})\,(T_A|_{\mathcal{R}(A^*)})^{-1}\,(T_B(\mathbf{x}')) = T_B(\mathbf{x}')$$

となっている．$P(\mathbf{x}) = \mathbf{x}'$ であるから

$$\begin{aligned} T_A\,(T_A|_{\mathcal{R}(A^*)})^{-1}\,T_B P(\mathbf{x}) &= T_A\,(T_A|_{\mathcal{R}(A^*)})^{-1}\,T_B(\mathbf{x}') \\ &= T_B(\mathbf{x}') = T_B(\mathbf{x}) \end{aligned}$$

が得られる．$(T_A|_{\mathcal{R}(A^*)})^{-1} T_B P$ の表現行列を $C$ とすると，$C$ が求めるものである．

**問題 11.1**　$\mathcal{R}(Q) = \boldsymbol{K}^N$ であるから，$\mathcal{R}(AQ) = \mathcal{R}(A)$.

**問題 13.1**　$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ の逆行列が存在し，$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ であるから $\mathcal{U}$ は基底である．

また，$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}$ の逆行列も存在し，$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ であるから $\mathcal{V}$ は基底である．したがって，

$$X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}, \quad X_{\mathcal{V}\leftarrow\mathcal{E}} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

したがって，

$$X_{\mathcal{V}\leftarrow\mathcal{U}} = X_{\mathcal{V}\leftarrow\mathcal{E}}X_{\mathcal{E}\leftarrow\mathcal{U}} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

ゆえに

$$\mathbf{x}_{\mathcal{V}} = X_{\mathcal{V}\leftarrow\mathcal{U}}\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \mathbf{x}_{\mathcal{E}} = X_{\mathcal{E}\leftarrow\mathcal{U}}\mathbf{x}_{\mathcal{U}} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

**問題 13.2**　$\Delta_A(\lambda) = \begin{vmatrix} 3-\lambda & 2 \\ -2 & -1-\lambda \end{vmatrix} = (1-\lambda)^2$ より 1 が固有値であるが，

$$A - I_2 = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

より $x_2 = s$, $x_1 = -s$ が解である．したがって，固有空間は

$$V_1 = \left\{ s \begin{pmatrix} 1 \\ -1 \end{pmatrix} : s \in \boldsymbol{K} \right\}$$

である．$\dim V_1 = 1$ より 1 の固有ベクトルで $\boldsymbol{K}^2$ の基底になるようなものはとれない．ゆえに $A$ は対角化できない．

$\Delta_B(\lambda) = \lambda(\lambda-1)(\lambda-3)$ であるから，$0, 1, 3$ が固有値である．

$$0 \text{ の固有ベクトルは } \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad 1 \text{ の固有ベクトルは } \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

$$3 \text{ の固有ベクトルは } \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

である．$P = \begin{pmatrix} -1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ とすると $P^{-1} = \begin{pmatrix} \frac{1}{3} & -1 & \frac{1}{3} \\ -1 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix}$ であり，

$$\begin{pmatrix} \frac{1}{3} & -1 & \frac{1}{3} \\ -1 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

**問題 13.3**　$A$ の固有値は $1, 2, 3$ である．1 の固有空間を求める．

$$\begin{pmatrix} \lambda - 2 & 0 & -1 \\ 0 & \lambda - 2 & 0 \\ -1 & 0 & \lambda - 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

より $x_3 = s$, $x_2 = 0$, $x_1 = -s$ であるから，

$$V_1 = \left\{ s \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} : s \in \boldsymbol{R} \right\}.$$

2 の固有空間を求める．

$$\begin{pmatrix} \lambda - 2 & 0 & -1 \\ 0 & \lambda - 2 & 0 \\ -1 & 0 & \lambda - 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

したがって，$x_3 = 0$, $x_2 = a$, $x_1 = 0$ であるから，

$$V_2 = \left\{ s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : s \in \boldsymbol{R} \right\}$$

である．3 の固有空間を求める．

$$\begin{pmatrix} \lambda - 2 & 0 & -1 \\ 0 & \lambda - 2 & 0 \\ -1 & 0 & \lambda - 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

より，$x_3 = a$, $x_2 = 0$, $x_1 = -a$ であるから，

$$V_3 = \left\{ s \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : s \in \boldsymbol{R} \right\}.$$

したがって，

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

による直交行列 $U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ により次のように直交化できる．

$$U^*AU = \begin{pmatrix} -\frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \\ 0 & 1 & 0 \\ \frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \end{pmatrix}\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

$B$ についても同様の計算で，

$$W = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

とすれば，これは直交行列で，

$$W^*BW = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

**問題 13.4**　$U, D$ は定理 13.26 の証明で用いたものとする．定理 13.27 の証明と同様に

$$\sup_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B \cap V} \langle A\mathbf{x}, \mathbf{x} \rangle = \sup_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle$$

が得られる．以下では $\displaystyle\sup_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle = \lambda_j$ を証明する．$V \in \mathcal{V}_{N-j+1}$ を任意にとる．$W = \mathrm{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\}$ とする．このとき，$V \cap W \neq \{\mathbf{0}\}$ である．なぜならもし $V \cap W = \{\mathbf{0}\}$ とすると，

$$\dim(V \oplus W) = N - j + 1 + j = N + 1$$

となってしまうからである．したがって，$\mathbf{x} \in B \cap V \cap W$ が存在する．このような任意

の $\mathbf{x}$ は

$$\mathbf{x} = \sum_{k=1}^{j} x_k \mathbf{e}_k, \quad \sum_{k=1}^{j} |x_k|^2 = 1$$

と表される．ゆえに

$$\langle D\mathbf{x}, \mathbf{x} \rangle = \sum_{k=1}^{j} \lambda_k |x_k|^2 \leq \lambda_j$$

が成り立つ．これより

$$\min_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle \leq \min_{\mathbf{x} \in B \cap V \cap W} \langle D\mathbf{x}, \mathbf{x} \rangle \leq \lambda_j$$

となっている．したがって，

$$\sup_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle \leq \lambda_j$$

である．いま特に $V_0 = \operatorname{span}\{\mathbf{e}_N, \cdots, \mathbf{e}_j\}$ とすると，$V_0 \in \mathcal{V}_{N-j+1}$ であり，任意の $\mathbf{x} \in B \cap V_0$ に対して，

$$\mathbf{x} = \sum_{k=j}^{N} x_k \mathbf{e}_k, \quad \sum_{k=j}^{N} |x_k|^2 = 1$$

と表せるので，

$$\langle D\mathbf{x}, \mathbf{x} \rangle = \sum_{k=j}^{N} \lambda_k |x_k|^2 \geq \lambda_j.$$

ゆえに

$$\lambda_j \geq \sup_{V \in \mathcal{V}_{N-j+1}} \min_{\mathbf{x} \in B \cap V} \langle D\mathbf{x}, \mathbf{x} \rangle \geq \min_{\mathbf{x} \in B \cap V_0} \langle D\mathbf{x}, \mathbf{x} \rangle \geq \lambda_j$$

が得られる．

**問題 13.5**　$S = \left\{\mathbf{x} \in \boldsymbol{K}^N : \|\mathbf{x}\| = 1\right\}$ は $\boldsymbol{K}^N$ の有界閉集合である．$f(\mathbf{x}) = \mathbf{x}^* A\mathbf{x}$ は $S$ 上の連続関数であるから，ある $\mathbf{z} \in S$ が存在し，任意の $\mathbf{x} \in S$ に対して，$f(\mathbf{x}) \geq f(\mathbf{z})\,(:= c) > 0$ が成立する．任意の $\mathbf{x} \in \boldsymbol{K}^N$, $\mathbf{x} \neq \mathbf{0}$ に対して，$\dfrac{\mathbf{x}}{\|\mathbf{x}\|} \in S$ であるから，

$$c \leq f\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) = \frac{f(\mathbf{x})}{\|x\|^2}.$$

**問題 13.6**　$\mathbf{y} = U^*\mathbf{x}$ とする．

$$\lambda_N = \langle A\mathbf{x}, \mathbf{x} \rangle = \langle D\mathbf{y}, \mathbf{y} \rangle = \sum_{j=1}^{N} \lambda_j |y_j|^2.$$

ゆえに

$$\lambda_N \left(|y_N|^2 - 1\right) + \sum_{j=2}^{N} \lambda_j |y_j|^2 = 0$$

である．ゆえに $|y_N| = 1$ でなければならない．$\|\mathbf{y}\| = 1$ より $\mathbf{y} = (0, \cdots, 0, y_N)^T$ であるから，$D\mathbf{y} = \lambda_N \mathbf{y}$ が得られる．したがって

$$A\mathbf{x} = UU^* AU\mathbf{y} = UD\mathbf{y} = \lambda_N \mathbf{x}.$$

**問題 15.1**　(1) ⇒ (2)　$A\mathbf{x} = \mathbf{b}$ とする．このとき

$$AX\mathbf{b} = AXA\mathbf{x} = A\mathbf{x} = \mathbf{b}.$$

(2) ⇒ (1)　$A = (\mathbf{a}_1, \cdots, \mathbf{a}_N)$ とする．このとき $\mathrm{rank}(A) = \mathrm{rank}(A, \mathbf{a}_j)$ であるから，方程式 $A\mathbf{x} = \mathbf{a}_j$ は可解である．したがって，$AX\mathbf{a}_j = \mathbf{a}_j$ が成り立っている．したがって $AXA = A$ を得る．

**問題 15.2**　$B^*B$ および $CC^*$ は $r$ 次正則行列である ($r = \mathrm{rank}(A)$)．したがって，$(B^*B)(CC^*) = B^*AC^*$ は正則行列である．また定理 15.8 より，

$$\begin{aligned} A^\dagger &= C^*(CC^*)^{-1}(B^*B)^{-1}B^* = C^*\left((B^*B)(CC^*)\right)^{-1} B^* \\ &= C^*\left(B^*AC^*\right)^{-1} B^* \end{aligned}$$

である．

**問題 15.3**　(1)　定義より $A$ は $A^\dagger$ の MP 一般逆行列であるから $(A^\dagger)^\dagger = A$ が得られる．

$A^*(A^\dagger)^*A^* = (A(A^\dagger)A)^* = A^*$，また $(A^\dagger)^*A^*(A^\dagger)^* = (A^\dagger AA^\dagger)^* = (A^\dagger)^*$，さらに $(A^*(A^\dagger)^*)^* = A^\dagger A = (A^\dagger A)^* = A^*(A^\dagger)^*$，最後に $((A^\dagger)^*A^*)^* = AA^\dagger = (AA^\dagger)^* = (A^\dagger)^*A^*$ が成り立つ．ゆえに $(A^\dagger)^* = (A^*)^\dagger$ が成り立つ．

**問題 15.4**　$f(x) = -\dfrac{1}{4} + x$ とすると，$f(1) = \dfrac{3}{4}$，$f(3) = \dfrac{11}{4}$，$f(4) = \dfrac{15}{4}$，$f(6) = \dfrac{23}{4}$ であり，$\overline{y} = \dfrac{1 + \dfrac{3}{2} + 5 + \dfrac{11}{2}}{4} = \dfrac{13}{4}$.

$$\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N = \begin{pmatrix} \frac{3}{4} \\ \frac{11}{4} \\ \frac{15}{4} \\ \frac{23}{4} \end{pmatrix} - \begin{pmatrix} \frac{13}{4} \\ \frac{13}{4} \\ \frac{13}{4} \\ \frac{13}{4} \end{pmatrix} = \begin{pmatrix} -\frac{5}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{5}{2} \end{pmatrix}$$

より，$\|\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_N\|^2 = 13$.

$$\mathbf{y} - \overline{y}\,\mathbf{1}_N = \begin{pmatrix} 1 \\ \frac{3}{2} \\ 5 \\ \frac{11}{2} \end{pmatrix} - \begin{pmatrix} \frac{13}{4} \\ \frac{13}{4} \\ \frac{13}{4} \\ \frac{13}{4} \end{pmatrix} = \begin{pmatrix} -\frac{9}{4} \\ -\frac{7}{4} \\ \frac{7}{4} \\ \frac{9}{4} \end{pmatrix}$$

より，$\|\mathbf{y} - \overline{y}\,\mathbf{1}_N\|^2 = \begin{pmatrix} -\frac{9}{4} & -\frac{7}{4} & \frac{7}{4} & \frac{9}{4} \end{pmatrix} \begin{pmatrix} -\frac{9}{4} \\ -\frac{7}{4} \\ \frac{7}{4} \\ \frac{9}{4} \end{pmatrix} = \frac{65}{4}.$

したがって，$R = \sqrt{\frac{4}{5}}$.

**問題 15.5**　$f(x) = 8 - 2x$ とする．

$f(1) = 6,\ f(2) = 4,\ f(3) = 2,\ f(4) = 0,\ f(5) = -2,\ \overline{y} = 2.$

$$\widehat{\mathbf{y}} - \overline{y}\,\mathbf{1}_5 = \begin{pmatrix} 4 \\ 2 \\ 0 \\ -2 \\ -4 \end{pmatrix},$$

$$\mathbf{y} - \overline{y}\,\mathbf{1}_5 = \begin{pmatrix} 8 \\ 3 \\ 0 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ -2 \\ -3 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 6 & 1 & -2 & -3 & -2 \end{pmatrix} \begin{pmatrix} 6 \\ 1 \\ -2 \\ -3 \\ -2 \end{pmatrix} = 54.$$

$R = \sqrt{\frac{40}{54}} = 0.86$

**問題 15.6**

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{pmatrix}, \quad X^T X = \begin{pmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix},$$

$$(X^TX)^{-1} = \begin{pmatrix} \frac{23}{5} & -\frac{33}{10} & \frac{1}{2} \\ -\frac{33}{10} & \frac{187}{70} & -\frac{3}{7} \\ \frac{1}{2} & -\frac{3}{7} & \frac{1}{14} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{pmatrix} \begin{pmatrix} 8 \\ 3 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ 13 \\ 13 \end{pmatrix},$$

$$\begin{pmatrix} \frac{23}{5} & -\frac{33}{10} & \frac{1}{2} \\ -\frac{33}{10} & \frac{187}{70} & -\frac{3}{7} \\ \frac{1}{2} & -\frac{3}{7} & \frac{1}{14} \end{pmatrix} \begin{pmatrix} 11 \\ 13 \\ 13 \end{pmatrix} = \begin{pmatrix} \frac{71}{5} \\ -\frac{50}{7} \\ \frac{6}{7} \end{pmatrix}.$$

したがって，求める 2 次関数は $f(x) = \frac{71}{5} - \frac{50}{7}x + \frac{6}{7}x^2$.

**問題 15.7**　$X = BAC$ とおく．$AXA = ABACA = ACA = A$，また

$$XAX = BACABAC = BABAC = BAC = X.$$

**問題 16.1**　$M \leq N$．$A^*A$ はユニタリー行列であるから，ある $N$ 次ユニタリー行列 $V$ が存在し，

$$V^*(A^*A)V = \begin{pmatrix} \sigma_1^2 & & O \\ & \ddots & \\ O & & \sigma_N^2 \end{pmatrix}$$

とできる．$V = (\mathbf{v}_1, \cdots \mathbf{v}_N)$ とすると，$\mathbf{v}_1, \cdots \mathbf{v}_N$ は $\boldsymbol{K}^N$ の正規直交基底である．$r = \mathrm{rank}(A^*A)$ とする．$r = \mathrm{rank}(A) \leq \min\{M, N\}$.

$r < M$ の場合，

$$\sigma_1^2 \geq \cdots \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_M^2 = 0$$

となっている．

$r = M$ の場合は，$\sigma_1^2 \geq \cdots \geq \sigma_r^2 = \sigma_M^2 > 0$ が成り立っている．

$$\mathbf{u}_j = \frac{1}{\sigma_j} A\mathbf{v}_j, \quad j = 1, \cdots, r$$

とおく．このとき $1 \leq i, j \leq r$ に対して，

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A\mathbf{v}_i, A\mathbf{v}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle \mathbf{v}_i, A^* A\mathbf{v}_j \rangle$$
$$= \frac{\sigma_j^2}{\sigma_i \sigma_j} \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \frac{\sigma_j}{\sigma_i} \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

より，$\mathbf{u}_1, \cdots, \mathbf{u}_r$ は $\boldsymbol{K}^M$ の正規直交系である．

もし，$r < M$ ならばある $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_M \in \boldsymbol{K}^M$ を加えて，$\mathbf{u}_1, \cdots, \mathbf{u}_M$ は $\boldsymbol{K}^M$ の正規直交基底となるようにできる．$U = (\mathbf{u}_1, \cdots, \mathbf{u}_M)$ はユニタリー行列である．さらに

$$U^* AV = U^* A (\mathbf{v}_1, \cdots \mathbf{v}_N)$$

$$= \begin{pmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_r^* \\ \mathbf{u}_{r+1}^* \\ \vdots \\ \mathbf{u}_N^* \end{pmatrix} (A\mathbf{v}_1, \cdots, A\mathbf{v}_r, A\mathbf{v}_{r+1}, \cdots, A\mathbf{v}_N)$$

$$= \begin{pmatrix} \langle A\mathbf{v}_1, \mathbf{u}_1 \rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_1 \rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_1 \rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_1 \rangle \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \langle A\mathbf{v}_1, \mathbf{u}_r \rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_r \rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_r \rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_r \rangle \\ \langle A\mathbf{v}_1, \mathbf{u}_{r+1} \rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_{r+1} \rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_{r+1} \rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_{r+1} \rangle \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \langle A\mathbf{v}_1, \mathbf{u}_M \rangle & \cdots & \langle A\mathbf{v}_r, \mathbf{u}_M \rangle & \langle A\mathbf{v}_{r+1}, \mathbf{u}_M \rangle & \cdots & \langle A\mathbf{v}_N, \mathbf{u}_M \rangle \end{pmatrix}$$

となっている．

ここで，$1 \leq i, j \leq r$ のときは，$\langle A\mathbf{v}_j, \mathbf{u}_i \rangle = \sigma_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \sigma_j \delta_{ij}$ である．

また $r+1 \leq i \leq M$, $1 \leq j \leq r$ の場合は，$\langle A\mathbf{v}_j, \mathbf{u}_i \rangle = \sigma_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$.

さらに $r+1 \leq j \leq N$ のとき，

$$\|A\mathbf{v}_j\|^2 = \langle A\mathbf{v}_j, A\mathbf{v}_j \rangle = \langle \mathbf{v}_j, A^* A\mathbf{v}_j \rangle = \sigma_j^2 \langle \mathbf{v}_j, \mathbf{v}_j \rangle = 0$$

となっている．これより (16.1) が得られる．

**問題 16.2** (1)

$$A^* A = \begin{pmatrix} 4 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 4 & 0 & 0 & 4 \end{pmatrix}$$

であるから，$\det(zI - A^*A) = z^2(z-2)(z-8)$．ゆえに固有値は $8, 2, 0, 0$ である．8 の固有空間 $V_8$ の正規直交基底を求める．

$$\begin{pmatrix} 4 & 0 & 0 & -4 \\ 0 & 7 & -1 & 0 \\ 0 & -1 & 7 & 0 \\ -4 & 0 & 0 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -\frac{1}{7} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

ゆえに $x_4 = s$, $x_3 = 0$, $x_2 = 0$, $x_1 = s$ が $(8I - A)\mathbf{x} = \mathbf{0}$ の解である．したがって，求めるベクトルは

$$\mathbf{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

である．$V_2$ の正規直交基底を求める．

$$\begin{pmatrix} -2 & 0 & 0 & -2 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -4 & 0 & 0 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

より，$x_4 = 0$, $x_3 = s$, $x_2 = s$, $x_1 = 0$ が $(2I - A)\mathbf{x} = \mathbf{0}$ の解である．したがって，

$$\mathbf{v}_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

が $V_2$ の正規直交基底である．$V_0$ の正規直交基底を求める．

$$\begin{pmatrix} 4 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 4 & 0 & 0 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

より，$x_4 = t$, $x_3 = s$, $x_2 = -s$, $x_1 = -t$ が $A\mathbf{x} = \mathbf{0}$ の解．したがって，

$$\mathbf{v}_3 = \begin{pmatrix} 0 \\ -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} -\dfrac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \dfrac{1}{\sqrt{2}} \end{pmatrix}$$

が $V_0$ の正規直交基底になっている．次に

$$\mathbf{u}_1 = \frac{1}{\sqrt{8}} A\mathbf{v}_1 = \begin{pmatrix} \dfrac{1}{2} \\ \dfrac{1}{2} \\ 0 \\ \dfrac{1}{2} \\ \dfrac{1}{2} \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}} A\mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

とする．$\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5$ を $\mathbf{u}_1, \cdots, \mathbf{u}_5$ が $\boldsymbol{K}^5$ の正規直交基底になるように選ぶ．たとえば

$$\mathbf{u}_3 = \begin{pmatrix} \dfrac{1}{2} \\ -\dfrac{1}{2} \\ 0 \\ \dfrac{1}{2} \\ -\dfrac{1}{2} \end{pmatrix}, \quad \mathbf{u}_4 = \begin{pmatrix} -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{pmatrix}$$

がそのようなものである．$U = (\mathbf{u}_1, \cdots, \mathbf{u}_5)$ とする．このとき，

$$U^*AV = \begin{pmatrix} 2\sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

である．また特異値分解は

$$A = 2\sqrt{2}\mathbf{u}_1\mathbf{v}_1^* + \sqrt{2}\mathbf{u}_2\mathbf{v}_2^*$$

であるが，ここで，

$$2\sqrt{2}\mathbf{u}_1\mathbf{v}_1^* = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \sqrt{2}\mathbf{u}_2\mathbf{v}_2^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

である．

(2)　$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$, $A^* = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix}$ とすると

$$A^*A = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix} \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}.$$

$\begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$ の固有ベクトルと固有値は

$$\left\{\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}\right\} \longleftrightarrow 0, \quad \left\{\begin{pmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{pmatrix}\right\} \longleftrightarrow 90, \quad \left\{\begin{pmatrix} \frac{1}{2} \\ 1 \\ 1 \end{pmatrix}\right\} \longleftrightarrow 360.$$

ゆえに特異値は $\sigma_1 = \sqrt{360} = 6\sqrt{10}$, $\sigma_2 = \sqrt{90} = 3\sqrt{10}$, $\sigma_3 = 0$ である．上記の固有ベクトルを長さが 1 になるように定数をかけると次のようになる．

$$\mathbf{v}_1 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}.$$

これより

$$V = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

とすると，$V$ は $A^*A$ を対角化するユニタリー行列である．そこで

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{6\sqrt{10}}\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix},$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \frac{1}{3\sqrt{10}}\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}\begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix}.$$

$U = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix}$ とすると,

$$U^*AV = \begin{pmatrix} \frac{3}{10}\sqrt{10} & \frac{1}{10}\sqrt{10} \\ \frac{1}{10}\sqrt{10} & -\frac{3}{10}\sqrt{10} \end{pmatrix}\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}.$$

これが特異値標準化である．また特異値分解は

$$A = 6\sqrt{10}\,\mathbf{u}_1\mathbf{v}_1^* + 3\sqrt{10}\,\mathbf{u}_2\mathbf{v}_2^* = \begin{pmatrix} 6 & 12 & 12 \\ 2 & 4 & 4 \end{pmatrix} + \begin{pmatrix} -2 & -1 & 2 \\ 6 & 3 & -6 \end{pmatrix}.$$

実際,

$$6\sqrt{10}\,\mathbf{u}_1\mathbf{v}_1^* = 6\sqrt{10}\begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 6 & 12 & 12 \\ 2 & 4 & 4 \end{pmatrix},$$

$$3\sqrt{10}\,\mathbf{u}_2\mathbf{v}_2^* = 3\sqrt{10}\begin{pmatrix} \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix}\begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} -2 & -1 & 2 \\ 6 & 3 & -6 \end{pmatrix}.$$

(3) $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$, $B = A^* = \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix}$ とし，これの特異値標準化と特異値分解を求める．

$$B^*B = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 2 & -4 & 4 \\ -4 & 8 & -8 \\ 4 & -8 & 8 \end{pmatrix}.$$

これの固有ベクトルと固有値は

$$\left\{ \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\} \longleftrightarrow 0, \quad \left\{ \begin{pmatrix} \frac{1}{2} \\ -1 \\ 1 \end{pmatrix} \right\} \longleftrightarrow 18$$

である．

$\sigma_1 = \sqrt{18} = 3\sqrt{2}$, $\sigma_2 = \sigma_3 = 0$ が特異値である．

18 の固有ベクトルの長さが 1 になるように $\mathbf{v}_1 = \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{pmatrix}$.

$B^*B$ の 0 の固有空間の正規直交基底として，たとえば次のものが選べる (グラム-シュミット).

$$\mathbf{v}_2 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix},$$

$$\mathbf{w}_3 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \left( \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right) \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \frac{4}{\sqrt{5}} \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 1 \end{pmatrix}.$$

$$\mathbf{v}_3 = \frac{\sqrt{5}}{3} \begin{pmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{2}{15}\sqrt{5} \\ \frac{4}{15}\sqrt{5} \\ \frac{1}{3}\sqrt{5} \end{pmatrix}.$$

$$V = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{15}\sqrt{5} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \\ \frac{2}{3} & 0 & \frac{1}{3}\sqrt{5} \end{pmatrix}$$ とおくと，これはユニタリー行列になっている．

また

$$\mathbf{u}_1 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

これに直交する長さ 1 のベクトルとして $\mathbf{u}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ をとり，

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

とする．すると，

$$U^*BV = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{15}\sqrt{5} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \\ \frac{2}{3} & 0 & \frac{1}{3}\sqrt{5} \end{pmatrix}$$

$$= \begin{pmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

したがって，

$$V^*AU = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & 0 \\ -\frac{2}{15}\sqrt{5} & \frac{4}{15}\sqrt{5} & \frac{1}{3}\sqrt{5} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -3\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

が求める特異値標準化である．特異値分解は，

$$A = 3\sqrt{2}\begin{pmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

**問題 16.3**　(1)　$\mu = \min_{\mathbf{x}\in \boldsymbol{K}^N} \|A\mathbf{x} - \mathbf{b}\|$ とおく．$\mathbf{x}, \mathbf{x}' \in X$, $0 \le t \le 1$ とする．このとき，

$$\begin{aligned}\mu &\le \|A(t\mathbf{x} + (1-t)\mathbf{x}') - \mathbf{b}\| = \|tA\mathbf{x} - t\mathbf{b} + (1-t)A\mathbf{x}' - (1-t)\mathbf{b}\| \\ &\le t\|A\mathbf{x} - \mathbf{b}\| + (1-t)\|A\mathbf{x}' - \mathbf{b}\| = \mu\end{aligned}$$

であるから $t\mathbf{x} + (1-t)\mathbf{x}' \in W$ である．

(2)　$\mathbf{x}, \mathbf{x}'$ をノルム最小の最小 2 乗解とする．$a = \|\mathbf{x}\|$, $b = \mathrm{Re}\langle \mathbf{x}, \mathbf{x}'\rangle$ とおく．$a = \|\mathbf{x}'\|$ である．$\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{x}' \in W$ であるから，

$$a \le \left\|\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{x}'\right\| \le \frac{1}{2}a + \frac{1}{2}a = a$$

である．したがって，

$$\begin{aligned}0 &= \left\|\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{x}'\right\|^2 - a^2 = \frac{1}{4}a^2 + \frac{1}{4}a^2 + \frac{1}{2}b - a^2 \\ &= \frac{1}{2}\left(b - a^2\right).\end{aligned}$$

したがって，$a^2 = b$ である．ゆえに $\|\mathbf{x} - \mathbf{x}'\|^2 = 2a^2 - 2b = 0$ が得られる．

**問題 17.1**　$Z = \widetilde{X} D_X^{-\frac{1}{2}}$ とおく．

$$\begin{aligned}Z_{\mathrm{ave}}(+, k) &= \frac{1}{N}\sum_{n=1}^{N} \frac{\widetilde{X}(n, k)}{\sqrt{S_X(k, k)}} \\ &= \frac{1}{\sqrt{S_X(k, k)}}\frac{1}{N}\sum_{n=1}^{N}(x_{nk} - X_{\mathrm{ave}}(+, k)) = 0.\end{aligned}$$

ゆえに $Z_{\mathrm{ave}} = \mathbf{0}$．これより $\widetilde{Z} = Z$ である．

$$S_Z(k, k) = \frac{1}{N}\langle Z(:, k), Z(:, k)\rangle = \frac{1}{NS_X(k, k)}\sum_{n=1}^{N}\left|\widetilde{X}(n, k)\right|^2 = 1.$$

**問題 17.2**　$\mathbf{a} = (a_j) \in \boldsymbol{R}^p$ とする．$Z = X\mathbf{a}$ より，$\widetilde{Z} = \widetilde{X}\mathbf{a}$ が得られる．したがって，

$$S_Z = \frac{1}{N}\widetilde{Z}^T\widetilde{Z} = \frac{1}{N}\mathbf{a}^T\widetilde{X}^T\widetilde{X}\mathbf{a} = \mathbf{a}^T S_X \mathbf{a}.$$

**問題 18.1**　定義より，$k = 0, \cdots, N-1$ に対して，

$$\mathcal{F}^{-1}(\mathbf{x})[k] = \frac{1}{N}\sum_{n=0}^{N-1} x[n]\,\omega^{nk} = \frac{1}{N}\mathcal{F}[\mathbf{x}][-k]$$

が成り立っている.

**問題 18.2**

$$\begin{aligned}\mathcal{F}^{-1}(\mathbf{x} *_c \mathbf{y})[k] &= \frac{1}{N}\mathcal{F}(\mathbf{x} *_c \mathbf{y})[-k] = \frac{1}{N}\mathcal{F}(\mathbf{x})[-k]\mathcal{F}(\mathbf{y})[-k] \\ &= N\mathcal{F}^{-1}(\mathbf{x})[k]\mathcal{F}^{-1}(\mathbf{y})[k].\end{aligned}$$

**問題 18.3** (1)

$$\begin{aligned}\mathcal{F}(\mathbf{x}_\bullet \mathbf{y}) &= \mathcal{F}\left(\mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}))_\bullet \mathcal{F}^{-1}(\mathcal{F}(\mathbf{y}))\right) = \frac{1}{N}\mathcal{F}\left(\mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) *_c \mathcal{F}(\mathbf{y}))\right) \\ &= \frac{1}{N}\mathcal{F}(\mathbf{x}) *_c \mathcal{F}(\mathbf{y}).\end{aligned}$$

(2)

$$\begin{aligned}\mathcal{F}^{-1}(\mathbf{x}_\bullet \mathbf{y}) &= \mathcal{F}^{-1}\left(\mathcal{F}\left(\mathcal{F}^{-1}(\mathbf{x})\right)_\bullet \mathcal{F}\left(\mathcal{F}^{-1}(\mathbf{y})\right)\right) = \mathcal{F}^{-1}\left(\mathcal{F}\left(\mathcal{F}^{-1}(\mathbf{x}) *_c \mathcal{F}^{-1}(\mathbf{y})\right)\right) \\ &= \mathcal{F}^{-1}(\mathbf{x}) *_c \mathcal{F}^{-1}(\mathbf{y}).\end{aligned}$$

**問題 18.4** $A = F^*\Delta_{\mathbf{m}}F$, $B = F^*\Delta_{\mathbf{m}'}F$ と表せる. ゆえに $AB = F^*\Delta_{\mathbf{m}}\Delta_{\mathbf{m}'}F = F^*\Delta_{\mathbf{m}_\bullet\mathbf{m}'}F$ である. ゆえに $AB$ は巡回行列である.

**問題 19.1**

$$\tau_1(x\bigstar\tau_{-1}y)[k] = x\bigstar\tau_{-1}y[k-1] = \sum_{n=0}^{N-1} x[n]\,y[k+n] = x\bigstar y[k],$$

$$\begin{aligned}\tau_1(x * \tau_1 y)[k] &= x * \tau_1 y[k-1] = \sum_{n=0}^{N-1} x[n]\,y[k-1-(n-1)] \\ &= x * y[k].\end{aligned}$$

**問題 19.2**

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix},$$

$$\widetilde{U} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{8\sqrt{2}} & 0 & \frac{1}{8\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{8\sqrt{2}} & 0 & -\frac{1}{8\sqrt{2}} \\ \frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} & -\frac{1}{8\sqrt{2}} & 0 \\ -\frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{8\sqrt{2}} & 0 \\ 0 & \frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} & -\frac{1}{8\sqrt{2}} \\ 0 & -\frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{8\sqrt{2}} \\ -\frac{1}{8\sqrt{2}} & 0 & \frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{8\sqrt{2}} & 0 & -\frac{1}{8\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

であるから，$U^*\widetilde{U} = I_4$ である．$\widetilde{U}$ の階数は 4 であり，$U$ の階数は 4 である．しかし $(U \quad \widetilde{U})$ の階数は 6 である．もし $\mathrm{span}\,\{\tau_{2j}\mathbf{u}\}_{j=0}^{3} = \mathrm{span}\,\{\tau_{2j}\tilde{\mathbf{u}}\}_{j=0}^{3}$ であれば，この行列の階数は 4 でなければならない．

**問題 19.3**

$$
\begin{aligned}
P_0\mathbf{x} &= U_0^*\widetilde{U}_0^*\mathbf{x} \\
&= \begin{pmatrix}
\tau_0 u_0[0] & \cdots & \tau_{2(M-1)}u_0[0] \\
\tau_0 u_0[1] & \cdots & \tau_{2(M-1)}u_0[1] \\
\vdots & \ddots & \vdots \\
\tau_0 u_0[N-1] & \cdots & \tau_{2(M-1)}u_0[N-1]
\end{pmatrix} \\
&\quad \times \begin{pmatrix}
\overline{\tau_0\widetilde{u}_0[0]} & \cdots & \overline{\tau_0\widetilde{u}_0[N-1]} \\
\overline{\tau_2\widetilde{u}_0[0]} & \cdots & \overline{\tau_2\widetilde{u}_0[N-1]} \\
\vdots & \ddots & \vdots \\
\overline{\tau_{2(M-1)}\widetilde{u}_0[0]} & \cdots & \overline{\tau_{2(M-1)}\widetilde{u}_0[N-1]}
\end{pmatrix}\mathbf{x} \\
&= \begin{pmatrix}
\sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[0]}\tau_{2j}u_0[0] & \cdots & \sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[N-1]}\tau_{2j}u_0[0] \\
\sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[0]}\tau_{2j}u_0[1] & \cdots & \sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[N-1]}\tau_{2j}u_0[1] \\
\vdots & \ddots & \vdots \\
\sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[0]}\tau_{2j}u_0[N-1] & \cdots & \sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[N-1]}\tau_{2j}u_0[N-1]
\end{pmatrix} \\
&\quad \times \begin{pmatrix}
x[0] \\
x[1] \\
\vdots \\
x[N-1]
\end{pmatrix}.
\end{aligned}
$$

したがって，

$$
\begin{aligned}
(P_0\mathbf{x})[k] &= \sum_{n=0}^{N-1}\sum_{j=0}^{M-1}\overline{\tau_{2j}\widetilde{u}_0[n]}\tau_{2j}u_0[k]x[n] = \sum_{j=0}^{M-1}\sum_{n=0}^{N-1}x[n]\overline{\tau_{2j}\widetilde{u}_0[n]}\tau_{2j}u_0[k] \\
&= \sum_{j=0}^{M-1}\left(\mathbf{x}, \tau_{2j}\widetilde{\mathbf{u}}_0\right)\tau_{2j}u_0[k].
\end{aligned}
$$

また

$$
\begin{aligned}
\widetilde{P}_0\mathbf{x} &= \widetilde{U}_0 U_0^* \mathbf{x}\\
&= \begin{pmatrix} \tau_0\widetilde{u}_0[0] & \cdots & \tau_{2(M-1)}\widetilde{u}_0[0] \\ \tau_0\widetilde{u}_0[1] & \cdots & \tau_{2(M-1)}\widetilde{u}_0[1] \\ \vdots & \ddots & \vdots \\ \tau_0\widetilde{u}_0[N-1] & \cdots & \tau_{2(M-1)}\widetilde{u}_0[N-1] \end{pmatrix}\\
&\quad \times \begin{pmatrix} \overline{\tau_0 u_0[0]} & \cdots & \overline{\tau_0 u_0[N-1]} \\ \overline{\tau_2 u_0[0]} & \cdots & \overline{\tau_2 u_0[N-1]} \\ \vdots & \ddots & \vdots \\ \overline{\tau_{2(M-1)} u_0[0]} & \cdots & \overline{\tau_{2(M-1)} u_0[N-1]} \end{pmatrix} \mathbf{x}.
\end{aligned}
$$

したがって，

$$
\begin{aligned}
\left(\widetilde{P}_0\mathbf{x}\right)[k] &= \sum_{n=0}^{N-1}\sum_{j=0}^{M-1} \overline{\tau_{2j}u_0[n]}\,\tau_{2j}\widetilde{u}_0[k]x[n] = \sum_{j=0}^{M-1}\sum_{n=0}^{N-1} x[n]\overline{\tau_{2j}u_0[n]}\,\tau_{2j}\widetilde{u}_0[k]\\
&= \sum_{j=0}^{M-1} (x, \tau_{2j}u_0)\, \tau_{2j}\widetilde{u}_0[k].
\end{aligned}
$$

残りの主張も同様に計算できる．

**問題 19.4**

$$
P_1^{(1)}\mathbf{x} =
$$

$$
\frac{1}{2}\begin{pmatrix}
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0\\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0\\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1\\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1
\end{pmatrix}
\begin{pmatrix} 0\\1\\2\\3\\4\\5\\6\\7 \end{pmatrix}
= \frac{1}{2}\begin{pmatrix} -1\\1\\-1\\1\\-1\\1\\-1\\1 \end{pmatrix},
$$

$$P_1^{(2)}\mathbf{x} =$$

$$\frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix},$$

$$P_1^{(3)}\mathbf{x} =$$

$$\frac{1}{8}\begin{pmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = 2\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

$$P_0^{(3)}\mathbf{x} = \frac{1}{8}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} = \frac{7}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

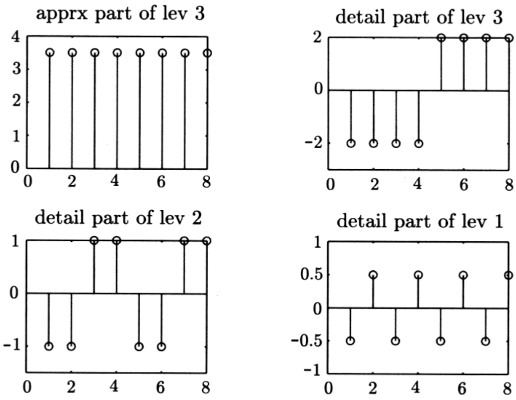



図 E. 1

**問題 19.5**

$$U_0^{(1)*} =$$

$$\frac{1}{4\sqrt{2}}\begin{pmatrix} 1+\sqrt{3} & 3+\sqrt{3} & 3-\sqrt{3} & 1-\sqrt{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1+\sqrt{3} & 3+\sqrt{3} & 3-\sqrt{3} & 1-\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1+\sqrt{3} & 3+\sqrt{3} & 3-\sqrt{3} & 1-\sqrt{3} \\ 3-\sqrt{3} & 1-\sqrt{3} & 0 & 0 & 0 & 0 & 1+\sqrt{3} & 3+\sqrt{3} \end{pmatrix},$$

$$U_0^{(1)} = \frac{1}{4\sqrt{2}}\begin{pmatrix} \sqrt{3}+1 & 0 & 0 & -\sqrt{3}+3 \\ \sqrt{3}+3 & 0 & 0 & -\sqrt{3}+1 \\ -\sqrt{3}+3 & \sqrt{3}+1 & 0 & 0 \\ -\sqrt{3}+1 & \sqrt{3}+3 & 0 & 0 \\ 0 & -\sqrt{3}+3 & \sqrt{3}+1 & 0 \\ 0 & -\sqrt{3}+1 & \sqrt{3}+3 & 0 \\ 0 & 0 & -\sqrt{3}+3 & \sqrt{3}+1 \\ 0 & 0 & -\sqrt{3}+1 & \sqrt{3}+3 \end{pmatrix},$$

$$U_1^{(1)*} =$$

$$\frac{1}{4\sqrt{2}}\begin{pmatrix} 1-\sqrt{3} & -3+\sqrt{3} & 3+\sqrt{3} & -1-\sqrt{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\sqrt{3} & -3+\sqrt{3} & 3+\sqrt{3} & -1-\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\sqrt{3} & -3+\sqrt{3} & 3+\sqrt{3} & -1-\sqrt{3} \\ 3+\sqrt{3} & -1-\sqrt{3} & 0 & 0 & 0 & 0 & 1-\sqrt{3} & -3+\sqrt{3} \end{pmatrix},$$

$$U_1^{(1)} = \frac{1}{4\sqrt{2}} \begin{pmatrix} -\sqrt{3}+1 & 0 & 0 & \sqrt{3}+3 \\ \sqrt{3}-3 & 0 & 0 & -\sqrt{3}-1 \\ \sqrt{3}+3 & -\sqrt{3}+1 & 0 & 0 \\ -\sqrt{3}-1 & \sqrt{3}-3 & 0 & 0 \\ 0 & \sqrt{3}+3 & -\sqrt{3}+1 & 0 \\ 0 & -\sqrt{3}-1 & \sqrt{3}-3 & 0 \\ 0 & 0 & \sqrt{3}+3 & -\sqrt{3}+1 \\ 0 & 0 & -\sqrt{3}-1 & \sqrt{3}-3 \end{pmatrix},$$

$$U_0^{(2)*} = \frac{1}{4\sqrt{2}} \begin{pmatrix} 1+\sqrt{3} & 3+\sqrt{3} & 3-\sqrt{3} & 1-\sqrt{3} \\ 3-\sqrt{3} & 1-\sqrt{3} & 1+\sqrt{3} & 3+\sqrt{3} \end{pmatrix},$$

$$U_0^{(2)} = \frac{1}{4\sqrt{2}} \begin{pmatrix} \sqrt{3}+1 & -\sqrt{3}+3 \\ \sqrt{3}+3 & -\sqrt{3}+1 \\ -\sqrt{3}+3 & \sqrt{3}+1 \\ -\sqrt{3}+1 & \sqrt{3}+3 \end{pmatrix},$$

$$U_1^{(2)*} = \frac{1}{4\sqrt{2}} \begin{pmatrix} 1-\sqrt{3} & -3+\sqrt{3} & 3+\sqrt{3} & -1-\sqrt{3} \\ 3+\sqrt{3} & -1-\sqrt{3} & 1-\sqrt{3} & -3+\sqrt{3} \end{pmatrix},$$

$$U_1^{(2)} = \begin{pmatrix} \sqrt{2}\left(-\frac{1}{8}\sqrt{3}+\frac{1}{8}\right) & \sqrt{2}\left(\frac{1}{8}\sqrt{3}+\frac{3}{8}\right) \\ \sqrt{2}\left(\frac{1}{8}\sqrt{3}-\frac{3}{8}\right) & \sqrt{2}\left(-\frac{1}{8}\sqrt{3}-\frac{1}{8}\right) \\ \sqrt{2}\left(\frac{1}{8}\sqrt{3}+\frac{3}{8}\right) & \sqrt{2}\left(-\frac{1}{8}\sqrt{3}+\frac{1}{8}\right) \\ \sqrt{2}\left(-\frac{1}{8}\sqrt{3}-\frac{1}{8}\right) & \sqrt{2}\left(\frac{1}{8}\sqrt{3}-\frac{3}{8}\right) \end{pmatrix}.$$

$$P_1^{(1)} = U_1^{(1)} U_1^{(1)*}, \quad P_1^{(2)} = U_0^{(1)} U_1^{(2)} U_1^{(2)*} U_0^{(1)*}, \quad P_0^{(2)} = U_0^{(1)} U_0^{(2)} U_0^{(2)*} U_0^{(1)*}$$

を計算すればよい．詳細は読者に委ねる．

**問題 21.3** $S_X = QQ^T$ をみたす正則行列 $Q \in \boldsymbol{R}^{p\times p}$ が存在する（たとえば定理 14.7 参照）．このとき $S_X^{-1} = \left(Q^T\right)^{-1} Q^{-1} = \left(Q^{-1}\right)^T Q^{-1}$ であるから，

$$d_{S_X}(\mathbf{x}, \mathbf{x}')^2 = \left\langle S_X^{-1}\mathbf{x}, \mathbf{x}' \right\rangle_{\boldsymbol{R}^p} = \left\langle \left(Q^{-1}\right)^T Q^{-1}\mathbf{x}, \mathbf{x}' \right\rangle_{\boldsymbol{R}^p} = \left\langle Q^{-1}\mathbf{x}, Q^{-1}\mathbf{x}' \right\rangle_{\boldsymbol{R}^p}.$$

ゆえに $T = Q^{-1}$ とすればよい．

**問題 22.1** 前半は明らかなので後半の主張を示す．$\langle\cdot,\cdot\rangle_{\boldsymbol{K}^M}$ を $\langle\cdot,\cdot\rangle_M$ というように表す．$\mathbf{x} = (x_j)$, $\mathbf{x}' = (x'_j) \in \boldsymbol{K}^M$, $\mathbf{y} = (y_j)$, $\mathbf{y}' = (y'_j) \in \boldsymbol{K}^N$ に対して，

$$\langle \mathbf{x}\otimes\mathbf{y},\mathbf{x}'\otimes\mathbf{y}'\rangle_{M\times N}=\sum_{j=1}^{M}\sum_{k=1}^{N}x_jy_k\overline{x'_jy'_k}=\sum_{j=1}^{M}x_j\overline{x'_j}\sum_{k=1}^{N}y_k\overline{y'_k}$$
$$=\langle \mathbf{x},\mathbf{x}'\rangle_M\langle \mathbf{y},\mathbf{y}'\rangle_N$$

が成り立っている．したがって，

$$\langle \mathbf{u}_j\otimes\mathbf{v}_k,\mathbf{u}_\mu\otimes\mathbf{v}_\nu\rangle_{M\times N}=\langle \mathbf{u}_j,\mathbf{u}_\mu\rangle_M\langle \mathbf{v}_k,\mathbf{v}_\nu\rangle_N=\begin{cases}1 & j=\mu,\,k=\nu\\ 0 & \text{その他}\end{cases}$$

であるから，$\{\mathbf{u}_j\otimes\mathbf{v}_k\}$ は正規直交系をなしている．基底であることは定理 22.5 (2) による．

**問題 22.3**　$\kappa\in L(X,Y;V)$ およびテンソル積の定義の (1) は明らかである．$\Phi\in L(X,Y;W)$ とする．$T\in L(V;W)$ を $T(\mathbf{e}_{jk})=\Phi(\mathbf{u}_j,\mathbf{v}_k)$ となるよう定義する．このとき，

$$T(\kappa(\mathbf{x},\mathbf{y}))=T\left(\sum_{j=1}^{M}\sum_{k=1}^{N}\alpha^j\beta^k\mathbf{e}_{jk}\right)=\sum_{j=1}^{M}\sum_{k=1}^{N}\alpha^j\beta^k\Phi(\mathbf{u}_j,\mathbf{v}_k)=\Phi(\mathbf{x},\mathbf{y}).$$

**問題 22.4**　$\mathbf{f}\in X^*$, $\mathbf{y}\in Y$ を任意にとり固定する．

$$\tau_{\mathbf{y},\mathbf{f}}(\mathbf{x})=\mathbf{f}(\mathbf{x})\mathbf{y},\quad \mathbf{x}\in X$$

と定めると $\tau_{\mathbf{y},\mathbf{f}}\in L(X;Y)$ である．また，

$$\tau(\mathbf{y},\mathbf{f})=\tau_{\mathbf{y},\mathbf{f}}$$

とすると $\tau$ は $Y\times X^*$ から $L(X;Y)$ への 2 重線形写像である．$\mathbf{x}_1,\cdots,\mathbf{x}_N$ を $X$ の基底で，その双対基底を $\mathbf{f}^1,\cdots,\mathbf{f}^N$ とし，$\mathbf{y}_1,\cdots,\mathbf{y}_M$ を $Y$ の基底とする．$\mathbf{e}_j^k=\tau(\mathbf{y}_j,\mathbf{f}^k)\in L(X,Y)$ とする．まず $\{\mathbf{e}_j^k\}_{j=1,\cdots,M,\,k=1,\cdots,N}$ が $L(X,Y)$ の基底になっていることを示す．$S\in L(X;Y)$ とする．

$$S(\mathbf{x})=\sum_{j=1}^{M}\alpha^j(\mathbf{x})\mathbf{y}_j$$

と表せる．$\mathbf{x}$ に $\alpha^j(\mathbf{x})$ を対応させる写像を $\alpha^j$ と表すと，これは $X$ 上の線形汎関数になっている．したがって，$\alpha^j=\sum_{k=1}^{N}\gamma_k^j\mathbf{f}^k$ と 1 次結合で表されている．したがって

$$S(\mathbf{x})=\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\mathbf{f}^k(\mathbf{x})\mathbf{y}_j=\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\tau_{\mathbf{y}_j,\mathbf{f}^k}(\mathbf{x})$$
$$=\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\mathbf{e}_j^k(\mathbf{x}).$$

また $\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\mathbf{e}_j^k=\mathbf{0}$ とすると，

$$0=\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\mathbf{e}_j^k(\mathbf{x}_l)=\sum_{j=1}^{M}\sum_{k=1}^{N}\gamma_k^j\mathbf{f}^k(\mathbf{x}_l)\,\mathbf{y}_j$$
$$=\sum_{j=1}^{M}\gamma_l^j\mathbf{y}_j$$

より $\gamma_l^j=0$ が成り立つ．以上より $\left\{\mathbf{e}_j^k\right\}_{j=1,\cdots,M,\,k=1,\cdots,N}$ が $L(X;Y)$ の基底であることがわかる．これより (1) が成り立つ．(2) を示す．$\Phi\in L(Y,X^*;W)$ に対して，

$$F\left(\mathbf{e}_j^k\right)=\Phi\left(\mathbf{y}_j,\mathbf{f}^k\right)$$

として $L(X;Y)$ から $W$ への線形写像 $F$ が定義できる．これは $F\left(\tau\left(\mathbf{y}_j,\mathbf{f}^k\right)\right)=\Phi\left(\mathbf{y}_j,\mathbf{f}^k\right)$ を意味しているので，このことから (2) が証明される．

**問題 22.5**　問題 22.4 の解答参照.

**問題 22.6**　記号の混乱をさけるために，この解答では，定義 22.3 の $T\otimes S$ を $T\widetilde{\otimes}S$ により表すことにする．$\left\{\mathbf{u}_j^{(i)}\right\}$ を $X_i$ の基底，$\left\{\mathbf{f}_j^{(i)}\right\}$ をその双対基底，$\left\{\mathbf{v}_j^{(i)}\right\}$ を $Y_i$ の基底とする $(i=1,2)$.

$$T_{i,j}(\mathbf{x})=\mathbf{f}_j^{(1)}(\mathbf{x})\,\mathbf{v}_j^{(1)},\quad S_{l,k}(\mathbf{x}')=\mathbf{f}_k^{(2)}(\mathbf{x}')\,\mathbf{v}_l^{(2)}$$

により $T_{i,j}\in L(X_1;Y_1)$, $S_{l,k}\in L(X_2;Y_2)$ を定義する．問題 22.5 (問題 22.4 の解答参照) より $T_{i,j}$ は $L(X_1;Y_1)$ の基底，また $S_{l,k}$ は $L(X_2;Y_2)$ の基底になっている．明らかに

$$\mathbf{f}_j^{(1)}\widetilde{\otimes}\mathbf{f}_k^{(1)}\in L(X_1\otimes X_2;\boldsymbol{K}\otimes\boldsymbol{K})=L(X_1\otimes X_2;\boldsymbol{K})=(X_1\otimes X_2)^*$$

であり，

$$\mathbf{f}_j^{(1)}\widetilde{\otimes}\mathbf{f}_k^{(1)}\left(\mathbf{u}_{j'}^{(1)}\otimes\mathbf{u}_{k'}^{(1)}\right)=\begin{cases}1 & j=j',k=k'\\ 0 & \text{その他}\end{cases}$$

であるから，$\left\{\mathbf{f}_j^{(1)}\widetilde{\otimes}\mathbf{f}_k^{(1)}\right\}$ は $\left\{\mathbf{u}_{j'}^{(1)}\otimes\mathbf{u}_{k'}^{(1)}\right\}$ の双対基底である．さらに

$$T_{i,j}\widetilde{\otimes}S_{l,k}(\mathbf{x}\otimes\mathbf{x}')=\mathbf{f}_j^{(1)}(\mathbf{x})\,\mathbf{f}_k^{(2)}(\mathbf{x}')\,\mathbf{v}_i^{(1)}\otimes\mathbf{v}_l^{(2)}=\mathbf{f}_j^{(1)}\widetilde{\otimes}\mathbf{f}_k^{(1)}(\mathbf{x}\otimes\mathbf{x}')\,\mathbf{v}_i^{(1)}\otimes\mathbf{v}_l^{(2)}$$

と問題 22.5 (問題 22.4 の解答参照) より，$T_{i,j}\widetilde{\otimes}S_{l,k}$ は $L(X_1\otimes X_2;Y_1\otimes Y_2)$ の基底になっている．したがって $\kappa$ は定義 22.1 (1) をみたす．

$W$ を線形空間で $\Phi\in L(L(X_1,Y_1),L(X_1,Y_1);W)$ とする．

$$F\left(\sum\alpha_{ijlk}T_{i,j}\widetilde{\otimes}S_{l,k}\right)=\sum\alpha_{ijlk}\Phi(T_{i,j},S_{i,j})$$

より $F\in L(L(X_1,Y_1)\otimes L(X_2,Y_2);W)$ を定義することができる．さらに

$$F(\kappa(T,S)) = F\left(T\widetilde{\otimes}S\right) = F\left(\sum \alpha\beta T_{i,j}\widetilde{\otimes}S_{l,k}\right) = \sum \alpha\beta\Phi\left(T_{i,j}, S_{i,j}\right) = \Phi(T,S)$$

が成り立つ．ゆえに $(L(X_1 \otimes X_2, Y_1 \otimes Y_2), \kappa)$ は $L(X_1; Y_1)$ と $L(X_2; Y_2)$ のテンソル積であることが示された．

**問題 22.7**　$A = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{M_1} \end{pmatrix}$, $B = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_{M_2} \end{pmatrix}$ とする．

$$T_A\left(\mathbf{e}_j^{(M_1)}\right) = \mathbf{a}_j = \sum_{\mu=1}^{N_1} a_j^\mu \mathbf{e}_\mu^{(N_1)},$$

$$T_B\left(\mathbf{e}_k^{(M_2)}\right) = \mathbf{b}_k = \sum_{\nu=1}^{N_2} b_k^\nu \mathbf{e}_\nu^{(N_2)}.$$

したがって，

$$T_A\left(\mathbf{e}_j^{(M_1)}\right) \otimes T_B\left(\mathbf{e}_k^{(M_2)}\right) = \sum_{\mu=1}^{N_1}\sum_{\nu=1}^{N_2} a_j^\mu b_k^\nu \mathbf{e}_\mu^{(N_1)} \otimes \mathbf{e}_\nu^{(N_2)}.$$

ゆえに，$\mathbf{z} = \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} z^{jk}\mathbf{e}_j^{(M_1)} \otimes \mathbf{e}_k^{(M_2)} \in \boldsymbol{K}^{M_1M_2}$ とすると

$$\begin{aligned}
T_A \otimes T_B(\mathbf{z}) &= \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} z^{jk} T_A \otimes T_B\left(\mathbf{e}_j^{(M_1)} \otimes \mathbf{e}_k^{(M_2)}\right) \\
&= \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} z^{jk} T_A\left(\mathbf{e}_j^{(M_1)}\right) \otimes T_B\left(\mathbf{e}_k^{(M_2)}\right) \\
&= \sum_{j=1}^{M_1}\sum_{k=1}^{M_2}\sum_{\mu=1}^{N_1}\sum_{\nu=1}^{N_2} z^{jk} a_j^\mu b_k^\nu \mathbf{e}_\mu^{(N_1)} \otimes \mathbf{e}_\nu^{(N_2)} \\
&= \sum_{\mu=1}^{N_1}\sum_{\nu=1}^{N_2}\left(\sum_{j=1}^{M_1}\sum_{k=1}^{M_2} z^{jk} a_j^\mu b_k^\nu\right) \mathbf{e}_\mu^{(N_1)} \otimes \mathbf{e}_\nu^{(N_2)}.
\end{aligned}$$

これより

$$T_C\begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \\ z^{21} \\ \vdots \\ z^{2M_2} \\ \vdots \\ z^{M_1M_2} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} a_j^1 b_k^1 z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} a_j^1 b_k^{N_2} z^{jk} \\ \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} a_j^2 b_k^1 z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} a_j^2 b_k^{N_2} z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1}\sum_{k=1}^{M_2} a_j^{N_1} b_k^{N_2} z^{jk} \end{pmatrix}$$

となっている．いま

$$a_1^1 B = \begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \end{pmatrix} = a_{11}\begin{pmatrix} b_1^1 & \cdots & b_{M_2}^1 \\ \vdots & \ddots & \vdots \\ b_1^{N_2} & \cdots & b_{M_2}^{N_2} \end{pmatrix}\begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{M_2} a_{11} b_k^1 z^{1k} \\ \vdots \\ \sum_{k=1}^{M_2} a_{11} b_k^{N_2} z^{1k} \end{pmatrix}$$

であるから次のことがわかる．

$$\begin{pmatrix} a_1^1 B & a_2^1 B & \cdots & a_{M_1}^1 B \end{pmatrix}\begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \\ z^{21} \\ \vdots \\ z^{2M_2} \\ \vdots \\ z^{M_1M_2} \end{pmatrix}$$

$$
= a_1^1 B \begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \end{pmatrix} + a_2^1 B \begin{pmatrix} z^{21} \\ \vdots \\ z^{2M_2} \end{pmatrix} + \cdots + a_{M_1}^1 \begin{pmatrix} z^{M_1 1} \\ \vdots \\ z^{M_1 M_2} \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{k=1}^{M_2} a_1^1 b_k^1 z^{1k} \\ \vdots \\ \sum_{k=1}^{M_2} a_1^1 b_k^{N_2} z^{1k} \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^{M_2} a_2^1 b_k^1 z^{2k} \\ \vdots \\ \sum_{k=1}^{M_2} a_2^1 b_k^{N_2} z^{2k} \end{pmatrix} + \cdots + \begin{pmatrix} \sum_{k=1}^{M_2} a_{M_1}^1 b_k^1 z^{M_1 k} \\ \vdots \\ \sum_{k=1}^{M_2} a_{M_1}^1 b_k^{N_2} z^{M_1 k} \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^1 b_k^1 z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^1 b_k^{N_2} z^{jk} \end{pmatrix}.
$$

したがって，

$$
\begin{pmatrix} a_1^1 B & a_2^1 B & \cdots & a_{M_1}^1 B \\ a_1^2 B & a_2^2 B & \cdots & a_{M_1}^2 B \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{N_1} B & a_2^{N_1} B & \cdots & a_{M_1}^{N_1} B \end{pmatrix} \begin{pmatrix} z^{11} \\ \vdots \\ z^{1M_2} \\ z^{21} \\ \vdots \\ z^{2M_2} \\ \vdots \\ z^{M_1 M_2} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^1 b_k^1 z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^1 b_k^{N_2} z^{jk} \\ \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^2 b_k^1 z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^2 b_k^{N_2} z^{jk} \\ \vdots \\ \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_j^{N_1} b_k^{N_2} z^{jk} \end{pmatrix}
$$

となり，

$$C=\begin{pmatrix} a_1^1B & a_2^1B & \cdots & a_{M_1}^1B \\ a_1^2B & a_2^2B & \cdots & a_{M_1}^2B \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{N_1}B & a_2^{N_1}B & \cdots & a_{M_1}^{N_1}B \end{pmatrix}$$

が導かれる．

**問題 22.9**　$X_j=X,\ j=1,2,\cdots$ とおく．$p\geq 2$ とする．

$$\Phi_p(\mathbf{x}_1,\cdots,\mathbf{x}_p)=S(\mathbf{x}_1)\cdots S(\mathbf{x}_p)$$

とすると，$\Phi_p\in L(X_1,\cdots,X_p;A)$ である．したがって，テンソル積の定義から，ある $T_p\in L(X_1\otimes\cdots\otimes X_p;A)$ で

$$\Phi_p(\mathbf{x}_1,\cdots,\mathbf{x}_p)=T_p(\mathbf{x}_1\otimes\cdots\otimes\mathbf{x}_p)$$

をみたすものがただ一つ存在する．$T_0(\alpha)=\alpha\mathbf{e}$, $\alpha\in\boldsymbol{K}=T^0(X)$ とし，また，$T_1=S$ である．$T(X)\ni\mathbf{x}=\sum\limits_{p=0}^{\infty}x_p$ に対して，

$$\widetilde{T}(\mathbf{x})=\sum_{p=0}^{\infty}T_p(\mathbf{x}_p)$$

とおくと $\widetilde{T}\in L(T(X);A)$ となる．いま $X$ の基底 $\mathbf{e}_1,\cdots,\mathbf{e}_N$ から定義される $T^p(X)$ の基底を $\mathbf{e}_{i_1\cdots i_p}$ とする．定義から $\mathbf{e}_{i_1\cdots i_p}$, $\mathbf{e}_{j_1\cdots j_q}$ に対して，

$$\widetilde{T}(\mathbf{e}_{i_1\cdots i_p}\otimes\mathbf{e}_{j_1\cdots j_q})=\widetilde{T}(\mathbf{e}_{i_1\cdots i_p})\widetilde{T}(\mathbf{e}_{j_1\cdots j_q})$$

である．これより容易に $\widetilde{T}$ が $T(X)$ から $A$ への準同形写像であることがわかる．$T_0$, $T_1$ の定義から (22.10) が得られる．定理の条件をみたす任意の写像 $\widetilde{S}$ に対して

$$\begin{aligned}\widetilde{S}(\mathbf{e}_{i_1\cdots i_p})&=\widetilde{S}(\mathbf{e}_{i_1})\cdots\widetilde{S}(\mathbf{e}_{i_p})=S(\mathbf{e}_{i_1})\cdots S(\mathbf{e}_{i_p})\\&=\widetilde{T}(\mathbf{e}_{i_1})\cdots\widetilde{T}(\mathbf{e}_{i_p})=\widetilde{T}(\mathbf{e}_{i_1\cdots i_p})\end{aligned}$$

より一意性も証明される．

**問題 A.1**　(1)

$$\begin{aligned}\sigma&=\begin{pmatrix}1&2&3&4\\2&4&1&3\end{pmatrix}=(1,2)\begin{pmatrix}1&2&3&4\\ \mathit{1}&4&\mathit{2}&3\end{pmatrix}\\&=(1,2)(2,4)\begin{pmatrix}1&2&3&4\\1&\mathit{2}&\mathit{4}&3\end{pmatrix}=(1,2)(2,4)(3,4).\end{aligned}$$

これを検算してみる．$(1,2)(2,4)(3,4)$ は

$$1 \to 1 \to 1 \to 2, \quad 2 \to 2 \to 4 \to 4$$
$$3 \to 4 \to 2 \to 1, \quad 4 \to 3 \to 3 \to 3.$$

(2)

$$\begin{aligned}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} &= (1,4)\begin{pmatrix} 1 & 2 & 3 & 4 \\ \mathit{1} & 3 & \mathit{4} & 2 \end{pmatrix} \\ &= (1,4)(2,3)\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & \mathit{2} & 4 & \mathit{3} \end{pmatrix} \\ &= (1,4)(2,3)(3,4).\end{aligned}$$

# 参考文献

本書を執筆する際に参考にした文献を記す．


[ABG] Adi Ben-Israel and T. N. E. Greville, Generalized Inverses, Theory and Applications, 2nd ed., Springer, 2003.

[新井 1] 新井仁之，フーリエ解析学，朝倉書店，2003.

[新井 2] 新井仁之，ウェーブレット，共立出版，準備中.

[CM] S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Dover Publ. Inc., 1979.

[C] O. Christensen, An Introduction to Frames and Riesz Bases, Birkhäuser, 2003.

[CDF] A.Cohen, I. Daubechies and J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, Comm. Pure and Appl. Math. (1992), 485–560.

[D] P. J. Davis, Circulant Matrices, John Wiley & Sons, 1979.

[F] M. W. Frazier, An Introduction to Wavelets through Linear Algebra, Springer, 1999.

[GMJ] M. Giaquinta, G. Modica and J. Soucek, Cartesian Currents in the Calculus of Variations, Vol.1 Cartesian Currents, Springer, 1998.

[GL] G. H. Golub and C. F. v. Loan, Matrix Computations, Johns Hopkins Univ. Press, 3rd ed., 1996.

[伊理] 伊理正夫，一般線形代数，岩波書店，2003.

[川久保] 川久保勝夫，線形代数学，日本評論社，1999.

[韓他] 韓太舜，伊理正夫，ジョルダン標準形，東京大学出版会，1982.

[L] D. C. Lay, Linear Algebra and its Applications, 3rd ed., Pearson Eduation, Inc., 2003.

[vL] C. F. van Loan, Generalizing the singular value decomposition, SIAM L. Numer. Anal., 13 (1976), 76–83.

[前原] 前原昭二，線形代数と特殊相対論，日本評論社，1981.

[MSY]J.H. マクラレン, R.W. シェーファー, M.A. ヨーダ, MATLAB による DSP 入門，ピアソン・エデュケーション，2000.

[M] D. Martin, Manifold Theory, An Introduction for Mathematical Physicists, Ellis Horwood, 1991.

[M] C. D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, 2000.

[野口] 野口潤次郎，複素解析概論，裳華房，1993.

[奥野他] 奥野忠一，久米均，芳賀敏郎，吉澤正，多変量解析法，日科技連，1971.

[PS] C. C. Paige - M. A. Saunders, Towards a generalized singular value decomposition, SIAM L. Numer. Anal., 18 (1981), 398–405.

[PW] D.B.Percival and A.T.Walden, Wavelet Methods for Time Series Analysis, Cambridge Univ. Press, 2000.

[佐武] 佐武一郎，線型代数学，裳華房，1958.

[佐武 2] 佐武一郎，線形代数，共立出版，1997.

[高木] 高木貞治，解析概論 (改訂第 3 版)，岩波書店，1961.

[砂田] 砂田利一，行列と行列式，岩波書店，2003.

[Sc] J. R. Schott, Matrix Analysis for Statistics, 2nd ed., Wiley, 2005.

[St] G. Strang, Linear Algebra and Its Applications, 3rd ed., Brooks/Cole, 1998.

[S] B. W. Suter, Multirate and Wavelet Signal Processing, Birkhäuser, 1998.

[竹村] 竹村彰通，統計，共立出版，1997.

[竹村・谷口] 竹村彰通，谷口正信，統計学の基礎 I，岩波書店，2003.

[富永] 富永晃，基礎演習，線形代数，聖文社，1975.

[内田・浦川] 内田伏一，浦川肇，線形代数概説，裳華房，2000.

[U] F. Uhlig, Transform Linear Algebra, Prentice Hall, 2002.

[Us] R. A. Usmani, Applied Linear Algebra, Marcel Dekker, 1987.

[V] P.Vaidyanathan, マルチレート信号処理とフィルタバンク，科学技術出版. 2001.

[柳井・竹内] 柳井晴夫，竹内啓，射影極限，一般逆行列，特異値分解，東京大学出版会，1983.

[横沼] 横沼健雄，テンソル代数と外積代数，岩波書店，1977.

[たのしみ] 数学のたのしみ，2004 秋 (日本評論社/亀書房) に所収のフォーラム『ウェーブレット解析の展開』.

# 索引


$r$ 次行列式因子　396
$*$ 作用素　473
アダマール積　27
アップサンプラー　364
アナログ信号　334

位相シフト　334
一意的に　260
1 次結合　9
1 次従属　138
1 次独立　138
一般解　194
一般化基本行列　35
一般化固有ベクトル　488
一般逆行列　279, 293
一般行列　273

$\mathbf{v}$ に関するハウスホルダー行列　36
$\mathbf{v}$ に関するハウスホルダー変換　36
ウェーブレット・フィルタ　358
上帯幅　267
上帯幅の帯行列　267
上三角行列　30
後ろ向き下三角行列　267
後ろ向き対角行列　267
後ろ向き対角成分　267

$A$ から定まる線形写像　25, 39
$A$ によって定義される線形写像　39
$A$ により定義される線形写像　25
$\mathbf{x}$ の基底 $\mathcal{U}$ による表示　416
$X$ の内積から誘導された $\Lambda^k(X)$ 上の内積　465
$\mathbf{x}$ の $\mathcal{U}$-座標　416
$N$ 次三角行列　30
$N$ 次基本行列　62
$N$ 次元基本ベクトル　9
$N$ 次元実ベクトル　4
$N$ 次元数ベクトル　4
$N$ 次元数ベクトル空間　7
$N$ 次元零ベクトル　9
$N$ 次元縦ベクトル　4
$N$ 次元複素ベクトル　4
$N$ 次元横ベクトル　9
$N$ 次下三角行列　30
$N$ 次巡回行列　115
$N$ 次正則行列　31
$N$ 次正方行列　18
$N$ 次対角行列　29
$N$ 次単位行列　28
$N$ 次置換　71
$N$ 次ファンデルモンド行列　106
$N$ 次フーリエ行列　339
$n$ 重線形形式　410
$N$ 点フーリエ変換　338
$N$ 点離散フーリエ変換　14
エネルギー　155
$(M,N)$ 上三角行列　30
$(M,N)$ 行列　18
$(M,N)$ 下三角行列　30
$(M,N)$ 対角行列　29
$M$ 行 $N$ 列の行列　18
$l^p$ ノルム　160
エルミート行列　230
エルミート形式　238

帯行列　267

解　47
回帰分析　324
回帰平方和　325
階数　180
階数分解　280
外積　450
ガウスの消去法　49
可解　193
核　44
角周波数　334

拡大係数行列　47
重ね合わせ　8
過剰フィルタ・バンク　367
完全再構成フィルタ・バンク　366
完全再構成条件　366

奇置換　77
基底　147, 411, 444
基底による表示　208
基本解　194
基本格子　399
基本平行体　402
逆行列　31
逆写像　43
逆写像をもつ　42
逆置換　74
$q$ 階共変テンソル　444
行階段数列　186
行既約階段行列　187
行に関する基本変形　57
行に関する初等変形　57
共分散　322
共変ベクトル　421
行列式　70, 79
行列式階数　393
寄与率　331
近似誤差　172
近似部分　378

偶置換　77
グラム行列　122
クロネッカー積　441
クロネッカーの記号　28
クロネッカーのデルタ記号　28

$k$-形式　455
$k$ 次小行列　118
$k$ 次小行列式　118
係数行列　47
$k$-ベクトル　455
$\boldsymbol{K}$ 上の多元環　448
$\boldsymbol{K}$ の多元環　449
ケーリー-ハミルトンの定理　488

広義の $n$ チャネル等間引きフィルタ・バンク　367
格子　399
合成写像　25
合成バンク　366
合成フィルタ　366
後退代入法　49, 55
恒等写像　32, 40
恒等置換　72
互換　74
個体　319
固有空間　221
固有多項式　117, 216
固有値　216
固有値問題　218
固有ベクトル　216
固有方程式　217
コレスキー下行列　264

最小多項式　488
最小 2 乗解　285
最小 2 乗型一般逆行列　293
最小 2 乗直線　289
最小 2 乗 $p$ 次曲線　292
最大間引きフィルタ・バンク　367
最良近似　172
差積　76
座標　208
座標系　207
サブサンプラー　363
サブバンド信号　376
作用素ノルム　311
三角不等式　423
残差　325
残差平方和　325
サンプリング　334
サンプリング間隔　334
サンプリング周期　334
サンプリング周波数　335
サンプル数　334
サンプル点　334

$(j,k)$ 成分　19
閾値　344
$\sigma$ に対する置換行列　112
次元　150, 412
辞書的配列　436
次数つき多元環　449
次数つき部分環　449
自然な埋め込み　415
自然な同形写像　415
下帯幅　267

下帯幅かつ上帯幅の帯行列　267
下帯幅の帯行列　267
下三角行列　30
実線形空間　408
実対称行列　230
実ベクトル空間　408
実歪対称行列　230
シフト・ベクトル　403
シフト変換　356
自明解　102
自明な解　102
自明な部分空間　145, 410
射影行列　171
斜交座標系　207
写像　10
シュヴァルツの不等式　423
重回帰直線　289
重回帰分析　289
重線形　409
重線形写像　409
重相関係数　291, 325
集中楕円　424
自由度　194
周波数　334
主成分得点　326
主成分ベクトル　326
循環相関積　345, 346
循環たたみ込み　341
準同形写像　449
詳細部分　378
商集合　402
商線形空間　411
ジョルダン・ブロック　487
ジョルダン標準形　487
信号　356
振幅　334
真部分空間　145

数ベクトル　4
スカラー　4
スカラー積　407
スカラー倍　8
スケーリング・フィルタ　358
スペクトル分解　238
スペクトログラム　354
スミス標準形　394

正規化周波数　335
正規行列　230
正規直交基底　160, 424
正規直交系　160
正規方程式　286
制限　153
斉次連立 1 次方程式　102
整数値行列　392
生成される部分空間　146, 410
正則　31
正則行列　31
正定値　241
正に向き付けられている　459
正の偶数　356
正の向きをもつ基底　459
成分　4
積　22, 72
絶対値　468
零行列　27
線形　11
線形演算　8
線形回帰　324
線形空間　407
線形空間の意味での同形　414
線形空間の意味での同形写像　414
線形形式　410
線形結合　9, 393, 410
線形写像である　11
線形従属　137, 393, 411
線形従属の関係にある　137
線形独立　138, 393, 411
線形汎関数　410
線形部分空間　145, 410
線形包　140
全射　43
前進代入法　56
全単射　43
全平方和　325

相関係数行列　322
相関係数　322
像空間　43, 44
双線形　410
双線形形式　410
双線形写像　410
双直交　371
双直交であるための必要十分条件　372
双対基底　413
双対空間　410

第 $i$ 行に関する余因子展開　97
第 $(i,j)$ 余因子　96
対角化可能　216
対角成分　30
第 $j$ 行，第 $k$ 列の成分　19
第 $j$ 成分　4
第 $j$ 列に関する余因子展開　97
互いに双直交　371
高さ $k$ の一般化固有ベクトル　488
多元環の意味で同形　449
多元環の意味での同形写像　449
多重解像度分解　378
たたみ込み作用素　341
縦ベクトル　4
単位インパルス　363
単位上三角行列　30
単位行列　28
単位元　449
単位下三角行列　30
単射　43

値域　43, 44
置換　71
チャネル $M_1$ 間引きフィルタ・バンク　367
直積線形空間　409
直和　152
直和分解　152
直交基底　161
直交行列　227
直交系　160
直交座標系　207
直交射影　167
直交射影行列　169
直交する　156
直交分解　165
直交補空間　166

低域通過フィルタ　344
ディジタル信号　334
データ行列　248, 320
データの標準化　324
テンソル積　6, 427, 435
テンソル積の結合律　434
テンソル積の双線形性　439
テンソル代数　450

同形　414, 428
同形写像　414, 428
同値関係　402
同値類　402
特異値　298
特異値標準化　300, 301
特異値標準形　300, 301
特殊解　194
特性多項式　117, 216
特性方程式　217
トレース　117

内積　155, 421, 423
内積から誘導される距離　423
内積から誘導されるノルム　423
内積空間　422
長さ　155

2 次形式　238

ノルム　155, 159, 423, 483
ノルム最小型一般逆行列　293

パラユニタリ　366
張られる部分空間　146
汎関数ノルム　425
反射型一般逆行列　293
半正定値　241
反変ベクトル　421

ピアソンの積率相関係数　322
$p$ 階交代テンソル　446
$p$ 階対称テンソル　446
$p$ 階反変 $q$ 階共変テンソル　444
$p$ 階反変テンソル　444
非自明解　102
非自明な解　102
非自明な部分空間　145, 410
非斉次連立 1 次方程式　102
非線形写像　12
左側から掛ける　23
左逆行列　277
非負定値　241
ピボット　186
ピボット選択　52
表現行列　39, 210
表現する行列　39
表示　210
標準基底　147
標準座標系　207
標準得点行列　324

標準偏差　322
ヒルベルト-シュミット・ノルム　483

フィルタ・バンク　366
フィルタリング　341
フーリエ係数　337
フーリエ・マルチプライヤ　343
複素線形空間　408
複素ベクトル空間　408
符号　77
負の向きをもつ基底　459
部分空間　145, 410
部分格子　401
部分格子を生成する　401
部分多元環　449
フローベニウス・ノルム　483
ブロック　125
ブロック行列　125
分解　251
分解可能　251
分解できる　260
分解バンク　366
分解フィルタ　366
分散　322
分散共分散行列　248, 321

平均ベクトル　246, 248, 321
平均偏差　321
平均偏差行列　321
平均偏差値　248
ベクトル空間　407
ベクトル積　475
変換　350
変換行列　211
変数　319
変量　319

ホッジ $*$ 作用素　473

マハラノビス距離　423

右側から掛ける　23
右逆行列　277

ムーア-ペンローズ一般行列　279
向きを定める　459
無限次元　412
無限直和　412

ヤコビアン　469

ユークリッド・ノルム　423
ユークリッド距離　423
ユークリッド内積　155, 422
有限次元　412
$\mathcal{U}$-座標写像　416
ユニタリー行列　227

横ベクトル　9
余積公式　470
予測値　325

ラグランジュ補間多項式　110
ランク　180

離散時間信号　334
離散フーリエ変換　14, 338
理想的フィルタ　344

累積寄与率　331

列に関する基本変形　493
列に関する初等変形　493
レベル 1 の多重解像度解析　380
レベル 2 の多重解像度解析　381
レベル $p$ の多重解像度解析　382
連続時間信号　334
連立 1 次方程式を解く　47

和　7, 407
歪エルミート行列　230

新井仁之（あらい・ひとし）

略歴
1959 年　横浜に生まれる．
1984 年　早稲田大学大学院理工学研究科修士課程修了．
現　在　東京大学大学院数理科学研究科教授．理学博士．
専攻は実解析学，調和解析学，ウェーブレット解析．

主な著書
『フーリエ解析と関数解析学』培風館
『ルベーグ積分講義』日本評論社
『フーリエ解析学』朝倉書店
『微分積分の世界』日本評論社

線形代数（せんけいだいすう）——基礎（きそ）と応用（おうよう）

2006 年 2 月 25 日　第 1 刷発行

著　者……………………新井仁之 
発行者……………………林　克行
発行所……………………株式会社 日本評論社
〒170-8474 東京都豊島区南大塚 3-12-4
電話：03-3987-8621［販売部］　http://www.nippyo.co.jp
企画・制作………………亀書房［代表：亀井哲治郎］
〒170-0005 東京都豊島区南大塚 3-12-3 ロダンサイド・ビル 4 階
電話 & FAX：03-3987-0248
E-mail：kame-shobo@nifty.com
印刷所……………………三美印刷株式会社
製本所……………………株式会社難波製本
装　幀……………………海保　透
ISBN 4-535-78519-8　Printed in Japan